

The University of Chicago

THE MEASUREMENT AND IMPROVE-
MENT OF SILENT READING AMONG
COLLEGE FRESHMEN

A PART OF A DISSERTATION SUBMITTED
TO THE FACULTY OF THE DIVISION OF THE
SOCIAL SCIENCES IN CANDIDACY FOR THE
DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF EDUCATION

By
IVAN ALBERT BOOKER

Private Edition, Distributed by
THE UNIVERSITY OF CHICAGO LIBRARIES
CHICAGO, ILLINOIS

1934

The University of Chicago

THE MEASUREMENT AND IMPROVE-
MENT OF SILENT READING AMONG
COLLEGE FRESHMEN

A PART OF A DISSERTATION SUBMITTED
TO THE FACULTY OF THE DIVISION OF THE
SOCIAL SCIENCES IN CANDIDACY FOR THE
DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF EDUCATION

By
IVAN ALBERT BOOKER

Private Edition, Distributed by
THE UNIVERSITY OF CHICAGO LIBRARIES
CHICAGO, ILLINOIS

1934



TABLE OF CONTENTS

LIST OF TABLES.	Page iv
LIST OF DIAGRAMS.	vii
LIST OF PLATES.	viii
Chapter	
I. INTRODUCTION	1
Statement of the Problem	
Significance of the Problem	
Scope of This Investigation	
Procedures Employed	
Organization of the Report	
II. DEVELOPMENT OF A BATTERY OF SILENT READING TESTS	5
The Need for Better Reading Tests	
Development of a Test of Rate of Continuous Reading	
Development of a Test of Reading Vocabulary	
Development of a Comprehension Test	
The Completed Battery of Silent Reading Tests	
III. THE READING ABILITY OF COLLEGE FRESHMEN	63
Typical Achievement as Revealed by Test Results	
Individual Differences as Shown by Test Scores	
Minimum Standards on the Reading Tests	
Relation between Rate and Comprehension	
Other Interrelationships	
Extent of Reading Difficulties	
The Significance of the Laboratory Study	
Description of Laboratory Procedure	
Presentation and Interpretation of the Data	
Summary	
IV. AN EXPERIMENT IN GROUP INSTRUCTION.	115
Relation of the Remedial Work to Other Phases of the Investigation	
Plans for the Remedial Work	
The Diagnosis of Reading Difficulties	
Objectives for the Work in Remedial Reading	
Procedure Followed in Remedial Instruction	
Results of the Experiment	
Evaluation of the Experiment	
V. INTERPRETATION.	166
Evaluation of the Findings Pertaining to Reading Ability	
Significance of the Work in Test Construction	
Evaluation of the Work in Remedial Reading	
Problems for Further Study	
BIBLIOGRAPHY.	179
APPENDIX.	184

LIST OF TABLES

Table	Page
I. The Reliability of Three Types of Rate Scores Based on Forms 1 and 2 of a Rate Test.	9
II. Correlation Coefficients between the Comprehension Scores of Various Groups on Forms 1 and 2 of a Rate Test.	11
III. Median Rate of Reading of Various Groups of Students on Forms 1 and 2 of a Rate Test.	12
IV. Median "Cr" Scores of Four Groups of Students on Two Forms of a Reading Test.	12
V. Difficulty of Each of the Comprehension Exercises of the Revised Rate Test	14
VI. Average Score of 35 College Freshmen on Comprehension Exercises Based on Test Paragraphs Which They Had Not Read.	22
VII. Per Cent of Difficulty and "V" for Each Comprehension Exercise of the "Test of Achievement in Silent Reading"	23
VIII. Reliability Coefficients for the "Test of Achievement in Silent Reading".	24
IX. Reliability Coefficients for the Comprehension Exercises of the "Test of Achievement in Silent Reading".	24
X. Median Scores of High-School Seniors on Four Forms of the "Test of Achievement in Silent Reading"	25
XII. Difficulty and Validity of the First Four Items of Form A	33
XIII. Supply of Usable Words Found in the Original Test Forms	34
XIV. Comparison of Four Forms of a Vocabulary Test with Respect to Difficulty and Validity.	34
XV. Median Scores of Two Groups of High-School Seniors on Five Forms of a 50-Word Vocabulary Test.	35
XVI. Number of Words at Each Level of Difficulty in the First 50-Word Vocabulary Tests and in the Revised Forms	37
XVII. Reliability Coefficients for Four Forms of the "Test of Reading Vocabulary"	38
XVIII. Comparison of the Mean and Standard Deviation for Four Forms of the "Test of Reading Vocabulary"	39
XIX. Validity of the Items of the "Test of Accurate Interpretation in Reading".	50
XX. Reliability of Various Parts of the "Test of Accurate Interpretation in Reading".	51

Table	Page
XXI. Reliability Coefficients for the "Test of Accurate Interpretation in Reading," Score "Cs"	53
XXII. Means for Four Forms of the "Test of Accurate Interpretation in Reading".	53
XXIII. Reliability of a Battery of Silent Reading Tests.	60
XXIV. Relative Difficulty of Two Forms of a Reading Test.	61
XXV. Distribution of the Scores of 664 Chicago Freshmen on Form B of the Vocabulary Test	66
XXVI. Distribution of the Comprehension Scores of 664 Chicago Freshmen.	66
XXVII. Correlation between Rate and Comprehension.	72
XXVIII. Correlation between Intelligence and Reading Ability.	74
XXIX. Correlation between Reading Test Scores and Number of Grade Points of 168 University Freshmen	76
XXX. Classification of 47 Students Who Withdrew from the University before the End of Two Quarters.	77
XXXI. Test Scores of Group I and Group II	85
XXXII. Summary of Eye-Movement Data Relating to Group I.	87
XXXIII. Summary of Eye-Movement Data Relating to Group II	88
XXXIV. Average Number of Fixations Per Line, Group I	90
XXXV. Average Number of Fixations Per Line, Group II.	91
XXXVI. Average Number of Regressive Movements Per Line, Group I.	92
XXXVII. Average Number of Regressive Movements Per Line, Group II	93
XXXVIII. Average Duration of Fixations in Twenty-Fifths of a Second, Group I	94
XXXIX. Average Duration of Fixations in Twenty-Fifths of a Second, Group II.	95
XL. Individual Performance Record of Student 14	96
XLI. Individual Performance Record of Student 56	96
XLII. A summary of Group Averages	103
XLIII. Correlation Coefficients between Certain Eye-Movement Measures.	110
XLIV. Eye-Movement and Intelligence Records of Ten Students in the Remedial Reading Class Compared with Those of Ten Other Students.	123

Table	Page
XLV. Data Concerning Students in the Remedial Reading Class.	125
XLVI. Achievement Records of Six Students Low in All Phases of Reading	127
XLVII. Achievement Records of Four Students Having Low Comprehension and Vocabulary Scores	132
XLVIII. Achievement Records of Four Slow Readers Having Satisfactory Comprehension	134
XLIX. Eye-Movement Records of Six College Freshmen on Two Paragraphs Supposedly of Equal Difficulty.	150
L. Initial and Final Reading Test Scores for Students in the Remedial Reading Class.	151
LI. The Eye-Movement Records of Sixteen Students before and after a Period of Training.	152
LII. Gains and Losses of Students in the Remedial Reading Class According to Their Final Test Scores and Eye-Movement Records . . .	153
LIII. Number of Fixations and Number of Regressive Movements per Line for Ten Pairs of College Freshmen in an Initial and Final Measurement	155
LIV. Per Cent of Students in the Remedial Class Who Gained in Various Phases of Reading	157
LV. Per Cent of Easy and Difficult Words Found in Sample Passages from Six Reference Books.	163
LVI. Number of Days' Absence for Students in the Remedial Class in Relation to the Progress Shown.	164

LIST OF DIAGRAMS

Diagram			Page
1. Classification of 457 Words with Respect to Their Relative Difficulty and Validity			33
2. Distribution of the Rate Scores of 664 Chicago Freshmen on Form 1 of the "Test of Achievement in Silent Reading"			64
3. Distribution of the Composite Scores of 664 Chicago Freshmen.			67

LIST OF PLATES

Plate			Page
I. Silent reading by five superior readers.			97
II. Silent reading by five poor readers.			98
III. Silent reading with few fixations per line			100
IV. Silent reading with many fixations per line.			100
V. Silent reading with few regressive movements per line.			101
VI. Silent reading with many regressive movements per line			101
VII. Silent reading with short fixation pauses.			102
VIII. Silent reading with long fixation pauses			102
IX. Silent reading typical of Groups I and II.			104
X. Silent reading by a student of average ability			104

ACKNOWLEDGMENTS

The writer acknowledges a debt of gratitude to all those who assisted with this investigation. The unfailing interest and timely counsel of Dr. William S. Gray have been particularly helpful. Special thanks is due also to Dr. Arthur E. Traxler for valuable assistance with many phases of the study. Dr. Floyd W. Reeves, Dean Chauncey S. Boucher, and Dr. Arthur P. Scott were responsible for administrative arrangements in the University of Chicago that made the study possible. A large number of high-school and college administrators, teachers, and students cooperated with the writer in developing the battery of reading tests, in obtaining the eye-movement records, or in the experiment in group instruction. In the preparation of the report, the writer has had the constructive criticism of a number of faculty members and students of the University of Chicago, as well as the timely suggestions and generous help of several members of the headquarters staff of the National Education Association. Finally, appreciation is expressed to Mrs. Booker, whose assistance with the statistical work and ready encouragement at all times have been invaluable.

Ivan Albert Booker

CHAPTER I

INTRODUCTION

Statement of the Problem

It is the purpose of this study to investigate the need for remedial reading among college freshmen and to evaluate one method of remedial instruction.¹ The investigation is concerned with the complete cycle in remedial work: discovery, diagnosis, and correction of reading difficulties. Pursuit of the problem implies (1) the selection or development, evaluation, and application of measurement techniques, (2) an analysis of reading abilities and disabilities, and (3) the development and evaluation of a program of remedial instruction. Considerable emphasis is placed on the development of techniques -- especially those of measurement -- for certain new techniques are employed. Nevertheless, the evaluation of the techniques and the analysis of the results obtained by applying them in a practical situation are more fundamental than their development. The study seeks to determine the general level of reading ability among college freshmen, the special abilities of superior readers, and especially the extent and character of reading deficiencies and the possibilities of correcting them through a particular type of group instruction.

Significance of the Problem

Since silent reading is recognized as an indispensable tool in study, it seems reasonable to expect that all college students should be able to employ it

1

The term "remedial" instruction is used in this study to refer to all the instructional procedures employed with a view to developing better reading habits. Those who are accustomed to distinguish between "remedial" teaching and "corrective" procedures would probably find the latter term more acceptable. However, since a broad interpretation of the word "remedial" is by no means uncommon, the word is so used in this report.

readily. Findings from a series of investigations beginning about 1920, however,¹ provide ample evidence that this is not the case. Research workers have found large groups of college students handicapped in reading. They have analyzed the difficulties repeatedly in so far as existing techniques would permit. They have studied the relationship between reading achievement and academic success in college. They have tried various types of remedial work.

In spite of the commendable work already done, the reading difficulties of college students still constitute an unsolved problem. Few satisfactory techniques for measuring reading achievement at the college level have been developed. Relatively little exact knowledge is available with reference to the silent reading habits of college students. Remedial procedures thus far reported seem far from satisfactory.

In view of these facts, further study of the reading achievement of college students with special reference to reading disabilities and their correction is urgently needed. Because the problem is so complex and because it varies from one institution to another, it must be attacked from many angles, by investigators working under different conditions, before colleges can apply research findings effectively in the interest of students handicapped in reading. The approach used in the present investigation is outlined below.

1

One of the early studies is reported in the following article:
C. W. Stone and Carl Colvin, "How to Study," Journal of Educational Psychology, XI (1920), 348-54.

Among the better-known studies that followed are:

W. F. Book, "How Well College Students Can Read," School and Society, XXVI (August, 1927), 242-48.

Mrs. L. C. Pressey, "College Remedial Reading Classes," English Journal, XIX (September, 1930), 566-69.

Alvin C. Burich, An Experimental Study of the Reading Abilities of College Students. Doctor's Dissertation, University of Minnesota, 1929.

Frank W. Parr, A Remedial Program for the Inefficient Silent Reader in College. Doctor's Dissertation, State University of Iowa, 1929.

Ernest M. Anderson, Individual Differences in the Reading Ability of College Students. University of Missouri Bulletin, XXIX (1928), Number 39.

Scope of This Investigation

The investigation reported here was focused on the silent reading habits of students in the 1930-31 freshman class of the University of Chicago. It sought to determine the extent and character of their reading difficulties and to evaluate a method of corrective instruction developed to meet their apparent needs.

Procedures Employed

In order to determine the reading difficulties experienced by Chicago freshmen and contribute in some degree to the ultimate elimination of this student handicap, three closely related steps were taken in the present investigation. The general nature of these steps is indicated in the brief description which¹ follows.

Test construction.- The investigation began with the construction of a battery of reading tests for use with college students, a step made necessary by the lack of appropriate measuring instruments.² The test battery comprised (1) a test of rate of continuous silent reading, (2) a vocabulary test, and (3) a comprehension test. Four forms were developed and partially standardized.

Measurement of reading achievement.- Measurement of the silent reading achievement of Chicago freshmen and an analysis of their silent reading habits

¹
Since the investigation required the use of several different methods of research, only a general statement of procedure is given here. A description of the method used accompanies the discussion of each phase of the study.

²
One of the serious obstacles to effective remedial work encountered by previous investigators was the lack of adequate measuring instruments. In many cases the tests employed were designed for younger students; in others, they yielded such incomplete or indefinite results that accurate diagnosis was impossible. No single battery of tests nor any combination of tests was found that constituted a satisfactory measuring instrument.

Other tests available were: The Iowa Silent Reading Test, the Minnesota Reading Test, the Whipple High School and College Reading Test, the Ingilis Test of English Vocabulary, the Markham English Vocabulary Tests, the Carr English Vocabulary Test, the Purdue Reading Tests, and the McClusky-Dolch Outline Test. The Nelson-Denny Reading Test appeared after this investigation was under way. An unpublished test has also been announced by Mrs. L. C. Pressey in the article: "College Students and Reading," Journal of Higher Education, II (January, 1931), 30-34.

constituted the second step in this investigation. The tests developed in this study and certain other tests were administered. Then in order to supplement the test results, eye-movement records were obtained for fifty students, twenty-five superior readers and twenty-five poor readers.¹ This phase of the study was undertaken in order to determine the need for remedial reading among Chicago freshmen, to identify students handicapped in reading, to provide a basis for the diagnosis of student difficulties, and to obtain additional information about the reading habits of mature students.

Remedial instruction.- The final step in the investigation was an experiment in remedial instruction. In view of the large number of students who were definitely handicapped in reading, an experiment in group instruction seemed advisable. Accordingly a plan was developed for organizing a remedial reading class in conjunction with a regular college course. Materials relating to this course were used as practice materials for the reading class. In other words, it was an attempt to develop better reading habits through the use of materials that students were expected to read in the preparation of regular assignments. Improvement in reading habits, as indicated from measurements obtained at the beginning and end of the training period, were used as a criterion in evaluating the experiment.

Organization of the Report

The organization of the report corresponds to the major steps outlined above. Chapter II deals with the development and evaluation of a battery of silent reading tests appropriate for college students. An analysis of silent reading ability based on test scores and eye-movement records is presented in Chapter III. A description and evaluation of the remedial work is given in Chapter IV. The final chapter interprets the data in terms of educational theory and practice and lists certain problems upon which further experimentation is needed.

1

The eye-movement records were obtained by means of the special camera in the laboratory of the School of Education, the University of Chicago. The procedure used is described in Chapter III.

CHAPTER II

DEVELOPMENT OF A BATTERY OF SILENT READING TESTS

The Need for Better Reading Tests

Proposed measurements.-- Only by the extensive use of reading tests could the reading habits and achievements of students in the 1930-31 freshman class of the University of Chicago be studied according to the plans outlined in Chapter I. Tests were to be used in four situations: (1) in making a general survey of the reading ability of the freshmen, (2) in retesting the students whose scores on the survey tests left any considerable doubt as to their abilities, (3) in determining the specific type or types of difficulty encountered by the students identified as poor readers, and (4) in measuring improvement in reading on the part of those who received remedial instruction.

General requirements for a suitable reading test.-- Certain general characteristics of a battery of reading tests appropriate for this study were apparent from the various uses enumerated above. For example, a suitable test would be one combining the advantages of a survey test with those of a diagnostic test. That is, a satisfactory test would be one with which poor readers could be identified promptly and one with which specific difficulties could be located. As stated above, it was to be used in both survey work and in diagnosis. It would have to be a test that could be administered to a large group of students at one sitting, inasmuch as it was to be given to more than six hundred students. Even in the follow-up work, economy of time and administrative convenience were recognized as important factors. Since repeated measurements were anticipated, an acceptable test would have three or four comparable forms. Obviously only a valid and reliable test would have any value in this study. It was further recognized that the test selected should be one that would give the remedial worker insight into the characteristic reading habits of college freshmen, provide him with dependable standards of achievement, and yield results which he could interpret in meaningful ways. Otherwise, the usefulness of the test in connection with the proposed remedial instruction would be distinctly limited. Finally, to serve the purposes of diagnosis mentioned above, the test

should measure whatever specific reading abilities are clearly distinguishable and essential in effective reading.

Abilities to be measured.¹- Three abilities are commonly accepted as basic factors in the total reading process: (1) ability to proceed rapidly, (2) ability to recognize word meanings, and (3) ability to grasp the meaning of the selection read. Although this simple classification is not exhaustive, few will deny that it is a useful analysis of the separable abilities that are fundamental to the reading process as a whole.

Accordingly, plans were laid for the development of a three-part reading test in four comparable forms. The construction of these tests will now be described. Although the rate, vocabulary, and comprehension tests are three parts of a battery rather than separate tests, they differ greatly from one another. For that reason the development of each part is discussed as though it were a separate test. Then, to give a more adequate picture of the test battery as a whole, the three parts are brought together for a brief evaluation in the final section of this chapter.

Development of a Test of Rate of Continuous Reading

The type of rate test that was needed.- In this investigation the most meaningful kind of rate measurement seemed to be one of normal rate of continuous silent reading. That is, in determining the reading ability of college students, the normal rate of silent reading seemed vastly more important than either rate of skimming, rate of rapid reading, or rate of careful, thoughtful reading. In the second place a rate score expressed in terms of "number of words per unit of time" seemed preferable to one expressed in terms of "point scores." The latter can be interpreted only in terms of relative performance on the test itself, whereas the former

1

As specific examples of the use of this classification see: (1) Paul V. Sangren, Improvement of Reading through the Use of Tests, Page 95. Kalamazoo, Michigan: Western State Teachers College, 1932; (2) Arthur E. Traxler, The Measurement and Improvement of Silent Reading at the Junior-High-School Level, pp. 19-35. Chicago: University of Chicago Libraries, 1932; and (3) Alvin C. Farich and M. E. Haggerty, Minnesota Reading Examination. Minneapolis: University of Minnesota Press, 1931.

describes the students' rate in definite meaningful terms.

A description of the trial forms.- In view of the needs stated above trial forms were developed in which the student was required to read a passage of continuous discourse. Narrative material was tried first but was discarded in favor of expository material of the essay type. Essays, twelve hundred words in length, on the topics "success" and "democracy" were used in the two forms developed first. These essays were written in simple style and were free from difficult or unusual words.

To encourage normal rate of reading the following precautions were taken:

(1) the phrases "rate test" and "speed test" were avoided in the title and directions; (2) the term "rate score" was also avoided; (3) neither rate nor comprehension was emphasized in the directions - the student was asked merely to read as rapidly as he could get the thought; and (4) the number of interruptions was reduced to the minimum in obtaining the rate score.

Since rate divorced from comprehension is meaningless, a check on comprehension was included in the test. This was entirely independent of the rate score and had no effect on rate except to suggest the quality of the reader's performance at the rate which he employed. Eight multiple-choice exercises were prepared to accompany each test form. The questions were so devised that the student's success would depend on his understanding of the author's meaning rather than on his memory of specific facts given.

In administering the tests, students were set to work with a starting signal. A measure of rate was obtained by having them "mark" the point reached in a given interval. All students were allowed to read the entire essay before answering the comprehension exercises based on it.

Questions involved in the preliminary work.- Trial use of the original test forms was expected to reveal the elements of strength and weakness in them: their reliability, their relative difficulty, the quality of the test exercises, and whether or not the essays were suitable for such a test. Moreover, rate may

be measured in either of two ways: (1) the amount read in a given time interval, or (2) the length of time required to read a given selection. The bearing of the experimental work on these problems is presented in the paragraphs which follow, using the scores obtained when the trial forms were administered to four different groups of students.

Method of administering and scoring the trial forms.- In this trial of the test forms, three measures of rate were obtained: (1) number of words read in 100 seconds, (2) number of words read in 200 seconds, and (3) rate for the selection as a whole. The first two measures were secured by having the student encircle the word he was reading at the end of 100 seconds and 200 seconds respectively. The total reading time was secured by having the student look up quickly as soon as he finished reading and copy the number displayed at that particular time.¹

The three scores, rate for 100 seconds, rate for 200 seconds, and average rate for the whole selection, were reduced to a common basis by converting them into words-per-second. The first rate score was obtained by pointing off two decimal places in the number of words read in 100 seconds. For the second, the number of words read in 200 seconds was divided by two and two places pointed off as before. For the third score, 1000 (the number of words in the essay) was divided by the several time intervals shown on the flash cards, a table was constructed of the various possible quotients, and a student's score obtained directly from the table. Since the last named method was unsatisfactory, it was abandoned.

Reliability of the rate score.- One of the problems upon which the preliminary work was focused was the probable reliability of such a measure of

1

One method of displaying the time indices was to use a set of flash cards, shifting the cards at intervals of five seconds. Another method was to record the time intervals on the blackboard every five seconds. The latter method was quite unsatisfactory. Aside from the error introduced by the examiner's failure to keep with the watch at all times, the method was distracting to many students. Furthermore, visual defects and poor lighting were difficult factors to control. Although the flash card method was much better, it was not entirely satisfactory. Both of these methods for securing total reading time are suggested by Stone in the manual of directions for his reading tests for grade-school children. C. R. Stone, Stone's Narrative Reading Tests, Bloomington, Ill.: Public School Publishing Co., 1923.

reading rate. The problem presented two issues in the present instance: (1) Which of the three contemplated types of rate measurement was most reliable? and (2) How reliable was the test in its original form? The reliability coefficients given in Table I suggested an answer to these questions.¹ From the data in this table one

TABLE I
THE RELIABILITY OF THREE TYPES OF RATE SCORES
BASED ON FORMS 1 AND 2 OF A RATE TEST

Group	Number of Cases	r ± P. E.		
		100"	200"	Total Time
Graduate students	13	.68 ± ?	-----	888 ± ?
Graduate students	27	.717 ± .068	-----	-----
High-school seniors	57	.669 ± .05	.766 ± .037	.715 ± .04
High-school seniors	84	.707 ± .037	.810 ± .025	.747 ± .03
Both college groups	40	.655 ± .06	-----	-----
Both high-school groups	141	.687 ± .03	.781 ± .022	.709 ± .028

fact seemed rather clear, namely, that a rate score based on the reader's performance for 100 seconds was less reliable than one obtained from a longer reading period.² Of the two other measures, the rate score obtained from the 200-second reading period showed the higher reliability. This was probably due to the greater accuracy of the measure rather than to the longer period of reading time

¹ The Pearson product-moment method of correlation is employed in all phases of this study unless otherwise stated, using a modification of the formula $r = \frac{\sum xy}{\sqrt{\sum x^2 \sum y^2}}$. See Karl J. Holzinger, Statistical Methods for Students in Education, p. 161. Boston: Ginn and Co., 1928.

² This conclusion is somewhat at variance with common practice in rate-test construction. Most rate tests of this type are based on an interval of one minute or less. The longest interval used in any test of this character is 100 seconds in the Gray Silent Reading Test. Sangren adopted the first minute of reading as the basis for the rate score on his test, but in the published report of his work (P. V. Sangren, Measurement of Achievement in Silent Reading, p. 16. Doctor's Dissertation, University of Michigan, 1927) he does not present data to show that this is the most reliable period. The conclusion presented here finds support, however, in the parallel work of Traxler on tests of a similar nature. (See: A. E. Traxler, op. cit., p. 22) In one instance, using the scores for 308 ninth-grade pupils, he found a reliability of .848 ± .011 for scores based on 200 seconds, and only .675 ± .021 for the scores of the same pupils on the same test forms when based on 100 seconds.

involved, although it is conceivable that wavering interest may have begun to impair the reliability of the test when performance during the total reading time was made the basis for the rate score.

The second phase of this analysis pertained to the adequacy of the proposed measure. Except in the case of the rate scores based on 100 seconds, all the reliability coefficients were higher than .70. All those based on the 200-second rate scores were higher than .75. In view of the fact that these were initial test forms the coefficients seemed gratifyingly high.

Evaluation of the score "Cr".-- The comprehension exercises were designed primarily to stabilize the rate score by discouraging superficial work and to indicate whether the student -- when reading at a given rate -- was able to understand the general point of view which was set forth in the essay. Experimental work with the test, however, suggested the possibility of making these exercises a part of the comprehension test. Many tests such as the Burgess, the Monroe, the Chapman-Cook, attempt to measure rate and comprehension simultaneously. No test, however, has attempted to measure comprehension by a series of exercises which depend upon the intelligent single reading of a body of continuous material. This ability is certainly one important phase of comprehension, especially in the study type of reading. No data can be presented here to establish the relative value of such a test of comprehension, nor is it implied that the comprehension exercises used with these tests are adequate measures of the ability which has been described. Nevertheless, the following results indicate why the idea of using these exercises as a part of the comprehension test was entertained.

The reliability coefficients for the comprehension exercises of the trial test forms are set forth in Table II. They were so low that these exercises were not considered as possible independent measures of comprehension, but, they were positive and significant. A test involving only eight items could scarcely be expected to be any more reliable. Of course, little significance could be attached to the first correlation in the table since it was based on an inadequate

sample, but the remainder of the coefficients seemed to lend plausibility to its relative size. The first three coefficients, based on results from advanced classes,

TABLE II
CORRELATION COEFFICIENTS BETWEEN THE COMPREHENSION SCORES
OF VARIOUS GROUPS ON FORMS 1 AND 2 OF A RATE TEST

Group	Number of Cases	$r \pm P. E.$
Graduate students	13	.78 $\pm$?
Graduate students	27	.519 $\pm$.09
Both college groups	40	.585 $\pm$.07
High-school seniors	57	.332 $\pm$.08
High-school seniors*	30	.461 $\pm$.10
High-school seniors	84	.346 $\pm$.06
Both high-school groups	141	.321 $\pm$.05

*

In the Chicago high school the test was administered to two sections or groups of seniors in English classrooms. This group was regarded by the English teacher as distinctly superior to the other.

were considerably higher than those obtained with high-school students. Furthermore, among the last four coefficients, the second, which was based on results from a superior section, was the highest. These facts suggested that the exercises were rather difficult for the average high-school senior; so much so, that chance probably operated to decrease the reliability of the test at that level. It was not expected that the test would continue to show a reliability of .78 or even .52 with advanced classes, but it seemed rather likely that the coefficients would be somewhat larger than .40. If that were true, the lower correlation with unselected classes of high-school seniors was not especially unfavorable. It merely indicated that the test was applicable with advanced classes in the college or university as well as with those at the entrance level. With the refinement of the test through the validation of the individual test items, reliability of .40 to .50 did not appear to be an unreasonable objective. In that case, the proposal that this score should be made a part of the total comprehension score seemed valid.

Relative difficulty of the test forms.— The median rate scores on the two forms of the test appear in Table III. With but two exceptions, the medians were higher for Form 1 than for Form 2, indicating that the latter was more difficult than the former.

TABLE III
MEDIAN RATE OF READING OF VARIOUS GROUPS OF STUDENTS
ON FORMS 1 AND 2 OF A RATE TEST

Group	Number of Cases	Adminis-tered First	100-second Median		200-second Median		Total Time Median	
		Form	1	2	1	2	1	2
Graduate students	13	1	4.00	3.57	---	---	3.46	3.67
Graduate students	27	2	3.72	3.90	---	---	---	---
Both college groups	40	—	3.83	3.79	---	---	---	---
High-school seniors	57	*	3.31	3.15	3.36	3.26	3.43	3.15
High-school seniors	84	1	3.72	3.25	3.52	3.32	3.60	3.52
Both high-school groups	141	—	3.53	3.21	3.46	3.36	3.54	3.38

* Form 1 was administered first to 27 students in one class section and Form 2 to the 30 students in another section.

The final comparison of the original test forms dealt with the relative difficulty of the comprehension exercises. The median "Cr" scores presented in Table IV revealed little actual difference in the difficulty of the two sets of

TABLE IV
MEDIAN "Cr" SCORES OF FOUR GROUPS OF STUDENTS ON
TWO FORMS OF A READING TEST

Group	Cases	Median on Form 1	Median on Form 2
Graduate students	13	6.6	6.5
Graduate students	27	6.6	6.6
High-school seniors	84	4.6	5.1
High-school seniors	57	5.2	4.4

exercises. They did show, however, that the exercises were hard enough to be discriminative. The criticism had been offered that an intelligent person could answer the questions correctly without having read the test material. These

medians fail to justify that inference. The exercises were difficult enough for these students, even after they had read the essays. Other data relating to this question will be presented in a later section of the chapter.

First revision of the test.- On the basis of the work just described the test forms were revised. The directions were modified slightly. In order to use a 200-second reading period, the essays were lengthened to 1,300 words. Two new comprehension exercises were designed for each of the test forms, and slight modifications were made in several of the original exercises. In a few instances they were made easier by the elimination of some phrase which had been too attractive as an alternate answer; for other exercises, more plausible decoys were formulated. Both forms were then administered to 114 high-school seniors and to 50 college freshmen and sophomores.

The median rate on Form 1, the "Essay on Success," was again considerably higher than that on Form 2. For the two groups tested the medians were 3.70 and 3.35, and 3.72 and 3.38. Since Form 1 was administered first to a part of each group and Form 2 to the remainder of each, the differences could not be attributed to practice effect. Form 2 was evidently the harder selection. The elimination of this difference was one objective of the next revision of the test. The median "Cr" scores indicated that the two comprehension parts of the tests were approximately equal in difficulty. With the larger group the median on Form 1 was 6.5 and on Form 2, 6.3. For the smaller group the medians were 7.4 and 7.5 respectively.

The rate score on the revised test showed approximately the same reliability as the original test forms. With the larger group (114 seniors) the coefficient was $.77 \pm .026$; with the smaller group it was $.72 \pm .049$. The reliability of the comprehension part of the test was somewhat higher than it had been with the original test forms. In the first instance the coefficients had been approximately .35. The coefficients in this case were $.55 \pm .045$ and $.43 \pm .077$, respectively. The increased reliability was probably due in part to the addition of the two new exercises and in part to the revision of the original ones.

To determine whether or not further improvement in the reliability of the "Cr" score might reasonably be expected, the difficulty of each test item was determined. A tally was made of the number of students (in the whole group of 164) who missed each item and the percentage of difficulty computed. The data are presented in Table V. The total number of items missed on Form 1 was practically identical to the number missed on Form 2, yet the individual test items showed wide differences in difficulty.

TABLE V
DIFFICULTY OF EACH OF THE COMPREHENSION EXERCISES
OF THE REVISED RATE TEST

Item Number	Number of Times Item Was Missed		Per Cent of Difficulty	
	Form 1	Form 2	Form 1	Form 2
1	49	26	30	16
2	86	37	52	23
3	30	25	18	15
4	31	108	19	66
5	67	142	41	87
6	78	55	48	34
7	122	24	74	15
8	63	124	38	76
9	64	72	39	44
10	57	33	35	20
Total	646	647	--	--

It is an accepted principle of test construction that items approximating¹ 50 per cent difficulty tend to be the better ones. Obviously, items passed by all students, or failed by all, have no discriminative value. Therefore, 50 per cent difficulty is one important criterion in evaluating single items.

¹
The soundness of this principle in test construction has been rather conclusively proved by Karl J. Holzinger and a number of his students who have worked on the problem. They have used test items from various fields, particularly Spelling and Arithmetic. Some experimental work in this connection was also done by the writer in connection with the evaluation of items for a vocabulary test. See especially Thelma G. Thurstone, The Relation Between a Test and Its Diagnostic Value. Doctor's Dissertation, University of Chicago Libraries, University of Chicago, 1926. Also, Clarence Sansom, The Diagnostic Value of a Test in Relation to Its Difficulty. Doctor's Dissertation, School of Education, the University of Chicago, 1929.

According to this standard, the exercises on Form 1 were the more satisfactory. Only two were easy for the class as a whole and only one was really hard. On Form 2, five exercises were rather easy and two were very hard. Therefore, it seemed likely that higher reliability might be secured if the exercises on Form 2 could be made more discriminative. A comparison such as this emphasizes the importance of subjecting every item of a test to separate study. Totals and averages give the test maker little indication of points at which the test is weak or strong.

Second revision of the test.- The two forms of the rate test were again revised. Four significant modifications were made in them. First, an introductory paragraph was written which could be used to introduce both essays. Thus, in all future work it was possible, without the slightest confusion, to have students sitting in adjoining seats taking different forms of the test. In eliminating practice effect this was an important advantage. Second, the essays were lengthened to 1,400 words in order that fewer students would finish reading within the 200-second period. Third, the sentence structure and vocabulary of Form 2 were simplified in a number of places in a deliberate attempt to make it easier. Finally, the comprehension exercises for Form 2 were revised. Since plans were made for an extensive trial of the material in this form, the test was printed instead of mimeographed. At the opening of the autumn quarter, 1930, Form 1 was administered to 318 entering freshmen at the University of Chicago and Form 2 to the remaining 325 students in the same class. Unfortunately for the purposes of test evaluation, both forms could not be given to all students. A little later in the autumn both forms were given to 168 college freshmen and 142 college sophomores.

The reliability of the rate score was much the same as that reported for earlier forms: $.737 \pm .023$ for the scores of the 168 freshmen, $.808 \pm .021$ for the scores of the 142 sophomores, and $.790 \pm .021$ for the two groups combined. The corresponding coefficients for the "Cr" score were: $.351 \pm .045$, $.516 \pm .043$, and $.436 \pm .031$. Note that the test showed higher reliability when used with sophomores than when used with freshmen. This was true with respect to both rate and comprehension. These data support an hypothesis based on the results obtained in the

earlier experimental work, namely, that the test is equally appropriate with advanced students.

The mean on Form 1 was consistently higher than that on Form 2 in the case of both rate and comprehension. The differences were not large; those reflected by the means for the two groups of Chicago freshmen are probably accurate indices of the real difference - rate, 3.86 and 3.50; "Cr", 7.2 and 6.5. The modifications intended to make Form 2 easier were not entirely successful. It was hoped that through further refinement the difference in difficulty indicated above could be reduced to some extent, but the possibility of equating rate tests exactly seemed rather remote.

Summary of the experimental work with rate test.- The foregoing results suggested that: (1) a rate test measuring the rate of uninterrupted continuous reading was feasible; (2) material of the essay type discussing some familiar abstract concept could be used satisfactorily at the college level; (3) the general technique employed in the experimental tests might well be continued in substantially the same form; (4) the rate score should be based on a reading period of 200 seconds in order to secure maximum reliability; (5) since the 200-second period seemed preferable and since many readers achieved a rate of more than six words per second, 1400-word passages should probably be used; (6) although the forms should be equated as nearly as possible with respect to rate, exact equality was unlikely to be attained; and (7) the comprehension exercises were reliable enough to constitute a partial measure of comprehension when carefully constructed and evaluated individually.

Procedure used in developing the final test forms.- Forms 1 and 2 had been modified so extensively both by revision and by the addition of new material that they bore only a slight resemblance to the original essays. It therefore seemed probable that additional essays prepared by the present writer would parallel those already tried out more closely than would material written by someone else. With that thought in mind another essay on "success" was prepared. The preparation of this one suggested the possibility for two others using the same topic.

These were also written. Each of the new essays was introduced by the same sentences used to introduce Form 1. They were written to parallel Form 1 as closely as possible with respect to organization, style, and general level of difficulty. Each contained 1400 words. Each dealt with two different factors that are important in the attainment of success, thus presenting the topic from different points of view.

The procedure which has been described in considerable detail in the foregoing sections was continued in the construction of the new test forms; namely, the development of a test, trial use of the material, revision of the form, trial, and revision to the point where the two forms were approximately equal in difficulty to Form 1. It was found to be very helpful to try the material first with individuals, and to keep an accurate record of the reading time paragraph-by-paragraph, both on the proposed test form and on Form 1. A number of graduate students in the School of Education, the University of Chicago, assisted in this phase of the work by reading the two forms in this manner and by criticizing the material. When the selections seemed satisfactory as far as could be determined by their use with individuals, the test forms were mimeographed and tried out with groups of high-school seniors or college students. Such further revisions were then made as seemed desirable, and the revised forms were tried again.

A step-by-step description of the development of these test forms and of the results secured in each trial use of the material would neither clarify the explanation of the method used, nor indicate the quality of the final test forms. By following the general procedure which has just been described three forms were made approximately equal in difficulty to Form 1, the development of which has been fully described. The original Form 2, entitled "An Essay on Democracy," was abandoned. One of these new essays, "Success," was substituted in its stead. The only change introduced into Form 1 was a very slight rewording of the directions in the interest of clearness and simplicity.

¹

A copy of the rate test Form 1 is presented in the Appendix, page 185.

The one technique used in the development of the final test forms, which has not been described already, was that of validating individual test items. A description of that procedure follows.

Validation of individual test items.— The assumption underlying this method is not a new one. It is that the best test item is the one to which all the good students (with respect to the ability in question) will give a correct response and to which all the poor students will give an incorrect response. Conversely, the poorest conceivable item is one which all the good students miss and all the poor students do correctly. An item of zero value is the one which is answered correctly by the same proportion of good and poor students. This assumption goes beyond the criterion of 50 per cent difficulty; an item might have exactly that degree of difficulty and still lack validity. For example, an ambiguous statement in the test exercise might account for 50 per cent difficulty and also for the fact that an equal number of both good and poor students got the item right. Many formulae have been proposed for determining the validity of an item. The most accurate evaluation is given by the method of biserial r , but the labor involved in employing that method makes it almost prohibitive. In fact, most of the proposed formulae involve such elaborate calculations that they can scarcely be employed in any major investigation. A very practical approach to the same objective is made possible by the following formula, which has been employed in the present investigation:

$$V = \frac{(R_u - W_u) - (R_l - W_l)}{N(u+1)}$$
$$\text{or } \frac{(R_u + W_l) - (W_u + R_l)}{N_u + 1}$$

In this formula "V" is the statistical constant which is being sought; "R" is

¹
This formula was suggested by Karl J. Holzinger and its value was studied as a class project under his direction. In that work a correlation of .58 was found between the values obtained by the use of this formula and those obtained by the method of biserial r . Clark's IV (Index of Validity) one of the widely known indices, showed a correlation of .50 with biserial r , and of .60 with V. For a description of the Index of Validity see, E. L. Clark, "Method of Evaluating the Units of a Test," Journal of Educational Psychology, XIX (April, 1928), 263.

the number of students who get an item right, and "W" the number who get it wrong. The letter "u" indicates a criterion group consisting of the highest 25 per cent of the whole group concerned, as measured by the score on the test as a whole or by some other measure of the ability in question. The letter "l" refers to a criterion group of the lowest 25 per cent of the students involved. " N_{u+l} " is the sum of the number of cases in the two criterion groups. The theoretical range of values yielded by the formula is +1, through 0, to -1, the same as that of the coefficient of correlation. Obviously, only those items with positive values would be considered, and the higher the value, the better the item.

In developing the final forms of the rate test, "V" as well as the percentage of difficulty was computed for each test item. Items that were poorest according to these criteria were carefully examined in an effort to discover why they were unsatisfactory and revised with a view to overcoming the observed defect

Description of the test in its present form.-- The rate test is entitled A Test of Achievement in Silent Reading. The first page provides information blanks, directions for the test, and spaces for the various scores. Since no part of the test material appears on this page, no student can get an unfair advantage of his classmates without turning the pages of the booklet, a violation of the directions which the examiner can easily note. The directions are relatively brief and simple. Page 2 contains ten multiple-choice comprehension exercises. They are printed on the page upside down so that the student sees them when he turns to that page but does not read them until he has finished reading the test essay. Below the comprehension exercises is the title of the test selection, "An Essay on Success," and the five lines used as a uniform introduction to each form of the test. The remainder of the 1400-word essay is printed on pages 3, 4, and 5 of the test booklet. Printed directions at the end of the essay and on page 2 with the comprehension exercises remind the student of what he is to do.

The test is administered in the following manner. After the students fill the blanks on the first page, the examiner reads the directions orally with them, asks them to turn to page 2, and directs their attention to the title of

the essay near the bottom of the page. He then reads orally the introductory part of the essay, the students following silently. When he comes to the last word on page 2, all turn to page 3, the examiner stops reading and notes the exact time, and the students continue to read silently, each at his own rate.¹ At the end of 100 seconds and again at the end of 200 seconds, each student marks the word which he is then reading and continues to the end of the essay. He then turns back and answers the comprehension exercises on page 2, indicating his answer to each question by placing a number in parentheses provided along the right-hand margin of the test blank.

Score A is a rate score based on the number of words read in 100 seconds; score B is a rate score representing half the number of words read in 200 seconds. With two decimal places pointed off, each score is expressed in terms of words-per-second. Score "Cr" is the number of exercises answered correctly.

Validity of the rate score.- To the extent that the rate score is reliable, it is valid by definition. It is the number of words read by the student in a given period of time. Few would maintain that the rate score is a measure of something other than rate. Two criteria may be used, however, to evaluate it on a basis other than that of mere definition: (1) the relative performance of the individual reader on material of different types, and (2) a comparison of rate as determined by the test with rate as determined by eye-movement records. Data pertaining to the first criterion are given in the following paragraph. Other data pertaining to both the first and the second criteria are presented in Chapter III, where the eye-movement records of 50 students are summarized.²

Does the slow reader on one type of material read slowly on other types of material and does the rapid reader in one instance read rapidly in another? In order to answer that question in part, Form 4 of the Test of Achievement in

1

The writer is indebted to W. S. Gray for the suggestion of the starting technique, and to A. E. Traxler for the placement of the comprehension exercises. The same starting technique was employed by J. E. Dransfield in an unpublished test developed by him at Columbia University, about 1925, but his work did not influence its selection for the tests developed in the present study.

2

See pp. 85-114.

1

Silent Reading and another rate test were administered to 46 college freshmen. The two rate tests were very similar in form, in purpose, and with respect to general mechanical features and plan of administration. The content, however, was very dissimilar. The former was a matter-of-fact exposition on success in life. The latter was a narrative involving adventure and describing a combat. The means reflect the difference in interest appeal; they were 3.73 and 4.23, respectively. In contrast to this clear evidence of difference in absolute performance, is that relating to relative performance, namely, the correlation between the two sets of scores. The coefficient was $.633 \pm .062$, which is only about one-sixth lower than the reliability of the tests themselves. Assuming that the tests had a reliability of .80, which is approximately correct, and applying Spearman's attenuation formula to correct the observed coefficient for response errors, the true correlation would be .79 instead of .633. This result may be somewhat exaggerated because of weakness in the attenuation formula. Nevertheless, it is suggestive of the true relationship. The essential point is that when two tests with a reliability of .75 to .80 show a correlation of .63 with each other the implied relationship is close. In this case it means that an individual's relative rate tends to remain fairly constant from one reading situation to another.

Validity of the comprehension exercises.- The validity of this part of the test rests in part on negative evidence. If the reader can give a correct response to a comprehension test question without having read the material on which it is based, the item is obviously worthless. In order to determine the extent to which students could reason out the correct answers to these exercises - the extent to

1

Form 1 of a rate test developed by A. E. Traxler.

2

See Holzinger, *op. cit.*, p. 253 for an explanation of the Spearman attenuation formula:

$$r_{st} = \frac{r_{12}}{\sqrt{r_{11} \cdot r_{22}}}.$$

3

Results obtained by A. E. Traxler (*op. cit.*, p. 20) confirm those presented above. He gave three informal reading tests to 92 high-school pupils. The content of the first was descriptive, the second narrative, and the third expository. His coefficients were: (1) descriptive with narrative, $r = .711 \pm .035$; (2) descriptive with expository, $r = .691 \pm .037$; (3) narrative with expository, $r = .729 \pm .033$.

which their replies were independent of the quality of their reading performance - the exercises were tested in the following manner.

A class of 35 college sophomores served as an experimental group. Page 2 of the Test of Achievement in Silent Reading was presented to each student. The class was then told to read each exercise carefully and select the most reasonable answer. Then to see what results could be expected from chance alone, they were told to fold under all the page except the exercise numbers and to put in the margin in front of each test item any digit between 1 and 5. They were urged to work as rapidly as possible and to "just put down any number between 1 and 5." The exercises for each of the four forms were evaluated in this way.

The results are summarized in Table VI. By chance alone, the student may be expected to make a score of 2. If the student has not read the paragraphs and makes an effort to reason out the answers his score is likely to be about 4. There is slight improvement with practice. It is apparent from these results that satisfactory performance on the exercises depends on the quality of the reading performance and not merely on the ability to reason.

TABLE VI

AVERAGE SCORE OF 35 COLLEGE FRESHMEN ON COMPREHENSION EXERCISES
BASED ON TEST PARAGRAPHS WHICH THEY HAD NOT READ

	Form 1	Form 2	Form 3	Form 4
Average from Chance	1.94	1.94	2.11	2.31
Average from Trial but without Reading the Test Paragraphs	3.20	3.54	3.91	4.57
Difference	1.26	1.60	1.80	2.26

Further evidence of the validity of this part of the test is the fact that the correlation between the rate score and the score "Cr" is very low. That is, this part of the test is not another measure of reading rate. Coefficients dealing with this relation are given in Chapter III, where the interrelationship of various abilities is discussed.

The validity of each test-exercise was computed, using the formula described on page 18. Each criterion group was composed of 30 high-school seniors selected on the basis of their performance on two forms of the test. The per cent of difficulty, or the per cent of students who got each item wrong, was also determined. The validity and difficulty of each exercise appear in Table VII. Every exercise has a positive value. The average "v" value is approximately .40. With the exception of

TABLE VII
PER CENT OF DIFFICULTY AND "v" FOR EACH COMPREHENSION EXERCISE
OF THE "TEST OF ACHIEVEMENT IN SILENT READING"

Question	Form 1		Form 2		Form 3		Form 4	
	Per Cent	v	Per Cent	v	Per Cent	v	Per Cent	v
1	.63	.67	.38	.53	.72	.27	.61	.13
2	.40	.53	.27	.27	.48	.27	.36	.47
3	.29	.27	.15	.27	.22	.33	.40	.73
4	.45	.40	.19	.27	.67	.20	.31	.47
5	.42	.13	.53	.33	.73	.53	.23	.67
6	.49	.27	.80	.13	.42	.53	.49	.53
7	.67	.60	.49	.60	.63	.73	.86	.07
8	.69	.07	.37	.27	.61	.27	.48	.47
9	.48	.67	.52	.53	.59	.33	.49	.73
10	.52	.80	.35	.27	.69	.47	.37	.87

six items the per cent of difficulty falls within the range of 25 to 75. Four items are below 25 per cent and two are above 75 per cent.

Reliability of the rate score.— Reliability coefficients presented in earlier sections of this chapter indicate in a general way the quality of the test with respect to this factor. When Forms 1 and 2 of the final rate test were administered to 59 high-school seniors in a Chicago high school, the reliability coefficient was $.81 \pm .03$. All forms of the test were then administered to 335 seniors in a large suburban high school. Each student took two forms of the test. Table VIII presents the reliability coefficients obtained in this instance and the number of cases on which each coefficient is based. The table should be read as follows: The correlation between the scores of the 60 students who took Forms 1 and 2 was $.800 \pm .03$; the coefficient from the scores of the 58 students who took Forms 1 and

3 was $.836 \pm .03$; and so on. Four of the coefficients are very good, one is fairly satisfactory, and one is somewhat low. Generally speaking, the reliability of the

TABLE VIII
RELIABILITY COEFFICIENTS FOR THE "TEST OF ACHIEVEMENT IN SILENT READING"

Form	$r \pm P. E.$			
	1	2	3	4
1	-----	$.800 \pm .03$	$.836 \pm .03$	$.762 \pm .04$
2	60 cases	-----	$.642 \pm .05$	$.827 \pm .03$
3	58 cases	57 cases	-----	$.594 \pm .06$
4	46 cases	58 cases	56 cases	-----

rate score approximates .75.

Reliability of the score "Cr."- The approximate reliability, or rather the probable range in the reliability of the score "Cr", is indicated by the coefficients in Table IX. These are based on the test papers of the 335 high-school seniors mentioned above and the arrangement of Table IX is identical to that of Table VIII.

TABLE IX
RELIABILITY COEFFICIENTS FOR THE COMPREHENSION EXERCISES
OF THE "TEST OF ACHIEVEMENT IN SILENT READING"

Form	$r \pm P. E.$			
	1	2	3	4
1	-----	$.446 \pm .07$	$.690 \pm .05$	$.571 \pm .07$
2	60 cases	-----	$.558 \pm .06$	$.721 \pm .04$
3	58 cases	57 cases	-----	$.603 \pm .06$
4	46 cases	58 cases	56 cases	-----

The coefficients are not consistent. It is highly probable that chance factors produced the high coefficient for Forms 2 and 4. However, five of the six coefficients fall above .50 and it seems unlikely that chance accounts for them all. The reliability of these exercises seems to be somewhat higher than .50. Forms 1 and 2 were used later with 157 college freshmen and the reliability in that instance was $.542 \pm .039$.

Relative difficulty of the final test forms.- When the final forms of the rate test were tried out with the 335 high-school seniors, in the testing program already mentioned, Form 1 was taken by 164 students, Form 2 by 175, Form 3 by 171, and Form 4 by 160. Any difference in rate due to practice effect was eliminated by giving each of the various forms in first and second position. The medians are summarized in Table X. The medians are not indetical, to be sure. Those for Forms 1 and 4 are very close. Those for Form 2 indicate that it is somewhat easier

TABLE X
MEDIAN SCORES OF HIGH-SCHOOL SENIORS ON FOUR FORMS OF
THE "TEST OF ACHIEVEMENT IN SILENT READING"

Form	1	2	3	4
Number of cases	164	175	171	160
Median rate	3.69	3.94	3.49	3.66
Median "Cr"	5.6	6.4	4.7	5.9

than any of the other forms with respect to both rate and comprehension.

The difficulty of equating two selections in an absolute manner, so far as rate is concerned, was demonstrated by the work with preliminary forms of the test. Differences in terms of test scores seem larger than they really are. For example, the difference between a rate of 4.00 and 3.60 is only 40 words or about four lines in 100 seconds. Because an individual's rate varies from day to day and even from one paragraph to another, chance factors may produce considerable variation in any set of scores. The same factors which keep reliability at approximately .75 make absolute equality between test forms difficult if not impossible. Although not standardized, these test forms seemed as nearly equal in difficulty as they could be made in this investigation. Further improvement would doubtless have been possible, but such refinements would have contributed little to the study as a whole in comparison with other projects which demanded attention.

Further evaluation of the test.- In the last section of this chapter and in various places in Chapters III and IV, facts are presented relative to the

merits and limitations of each part of the test battery. Therefore, the rate test should not be judged solely by the foregoing description, but in part by its usefulness in later phases of the study. The discussion now turns to the other units in the test series: the vocabulary test, and the comprehension test.

Development of a Test of Reading Vocabulary

A vocabulary test, like that in any other field, must be adapted to the particular purpose which it serves. That is, a test designed to measure the vocabulary used in oral expression is not appropriate for measuring the vocabulary used in writing, or that used in reading. Two criteria for the proposed vocabulary test were regarded as fundamental: (1) the test should measure ability to recognize word meanings, and (2) the score should indicate level of ability rather than speed of performance. The practical implications of these standards are pointed out in the following discussion of basic questions and problems.

How difficult should the words be?- Not all words are of equal importance in reading since some words are encountered much more frequently than others. To be uninformed as to the meaning of a very rare word is no particular handicap to the reader, but to be unfamiliar with a commonly used term may be a real obstacle to effective reading. Hence it is of utmost importance that a reading test should measure the extent of the reader's vocabulary with words of real social significance. On the other hand, because of their frequent occurrence, the commonly used terms tend to be learned more quickly than the others and it is acquaintance with the words used less frequently which really differentiates mature readers. One of the major problems in constructing a test of the reading vocabulary is therefore the maintenance of a proper balance between easy and difficult words. Both oral and written usage must be taken into account in considering the difficulty of a word. Mere frequency of occurrence in print is not an absolute criterion of difficulty, since many words that are rarely used in print are employed extensively in oral expression and are therefore quite generally known, and vice versa.

Test makers have faced the problem of securing words of proper difficulty in a variety of ways, but probably none has done it more effectively than Inglis in the construction of his Tests of English Vocabulary. His work was based on a series of assumptions which seem to be fundamentally sound,¹ although he offered no proof of their validity. Inglis designated three classes or types of words: (1) common words, those which are so frequently used that they are known by practically everyone; (2) very difficult words, those which are technical, rare, or so restricted in use that they are known by a relatively small number of persons; and (3) "the intelligent reader's vocabulary." Inglis described the latter as a vast body of words between the two extremes; a body of words of unquestioned social value, but known only in proportion to the breadth of one's intellectual interests and achievements. He drew all his test words from a "reservoir" which had been carefully evaluated by expert opinion and selected by the judges as words belonging in the "intelligent reader's vocabulary." This conception is in complete agreement with the principle that the best test items are neither too easy nor too hard - that 50 per cent difficulty is one appropriate criterion.² For a test designed for use with college students, words drawn from the "intelligent reader's vocabulary" seemed appropriate with respect to the general level of difficulty involved.

What constitutes an adequate sample?- All vocabulary tests must make use of the sampling method. Undoubtedly, the larger the sample the better the test will be. On the other hand, practical considerations always make it necessary to use a rather limited sample. What, then, is the minimum number of words which may be used and satisfactory reliability still be obtained?

Test makers have adopted different practices. Some tests contain not more than 40 words. Others have as many as 300 items. Inglis stated that in his

1

A report of his technique appears in an article entitled, "A Vocabulary Test for High-School and College Students," English Leaflet, XXIII (October, 1923), 1-13.

2

See p. 14.

experimental work, a test shorter than 150 words proved to be unreliable.¹ In view of the lack of agreement on this point, further study of the problem seemed desirable before reaching any decision as to the length of the proposed tests.

Preliminary work with vocabulary tests.- Since the Inglis test was based on a series of principles in harmony with those adopted in this investigation, an analysis of that test was made the point of departure in developing a new series of vocabulary tests. This work was focused primarily on the question: Can a reliable test shorter than the Inglis test be devised? A study was made of the relative difficulty of the individual items in the Inglis test;² the test was divided into three equal parts and the reliability of these parts was determined; and a study was made of the effect of scaling items according to difficulty, allowing equal intervals between succeeding items.

A vocabulary test devised for junior-high-school students was administered to 84 high-school seniors and to 27 graduate students.³ The words in this test had been drawn from the harder words in Thorndike's list of 10,000 words most commonly used. The test was given to the groups mentioned above in order to determine whether or not any words in Thorndike's list were difficult enough for a vocabulary test designed for use at the college level. The preliminary work just described seemed to warrant the following conclusions: (1) a reliable 50-word test could be devised, (2) words should be chosen which approximate 50 per cent difficulty as nearly as possible - the range probably should not exceed 25 to 75 per cent difficulty, (3) alphabetical arrangement of the words would be as satisfactory as any other, and (4) although a test based entirely on the Thorndike list was not difficult enough for college students, part of the words for the proposed tests could

¹
Inglis, op. cit.

²
Table XI in the complete manuscript of this Dissertation, on file in the Library of the University of Chicago, presents data relating to the difficulty of the words in the Inglis test.

³
This test was the Reading Vocabulary Test devised by Arthur E. Traxler.

be drawn from this source.

Appropriate mechanical features.- Other questions considered at the beginning of this project were: What is the most effective way in which to present the test words? What is the most desirable form of response? What is the best method of scoring? and, What is the proper place of this test in the series? Since the proposed vocabulary test was expected to measure the reading vocabulary, it appeared logical that each test word should appear in a short phrase, clause, or sentence¹ rather than in isolation. The multiple-choice technique was selected as an appropriate form of response, for if the reader cannot identify the synonym of a word, it is doubtful whether he would be able to give it a reasonably accurate interpretation in an actual reading situation. Furthermore, the multiple-choice technique is² objective and relatively free from the effects of guessing. Since the test was to be used to determine the quality of the reader's vocabulary, it seemed clear that it should be a power test. For that reason, it would have to be administered as an untimed test or a generous time allowance would have to be given to it. Final decision on this matter was made after trial forms had been used experimentally.

Preliminary word-lists for the new test forms.- After the length and general character of the proposed test had been thus determined, the next step was that of selecting test words. A supply of test words many times the length of the proposed tests would evidently be needed. To meet this need, the following procedure was adopted.

The words in the "upper five thousand" of Thorndike's list of 10,000 most commonly used words were copied in alphabetical order as one of the possible sources for usable words. As a supplement to this list, a random sample of the words in the dictionary was taken. The New International Dictionary was used in the latter

1

This plan is followed in the Carr English Vocabulary Test, the Inglis Test of English Vocabulary, the Markham English Vocabulary Test, and others.

2

This technique has been employed in virtually all recent vocabulary tests. For a theoretical discussion of its merits and limitations, see G. M. Ruch, op. cit., p. 274 ff.

project. The method of sampling was as follows:

1. The last defined word in each of the three columns on each page was copied except in the following instances:

- a. Proper names such as "Abel," "Abraham," etc., were rejected.
- b. Hyphenated words were taken, but double words such as "apple tree" were rejected.
- c. Foreign words, identified as such in the dictionary, were omitted; for example, the word "entrepreneur."

2. When the last defined word was a proper name, a double word, or a foreign word, the preceding word was taken.

3. Words with a technical meaning, indicated as such in the definition, and nouns referring to objects found in a limited geographical area were checked with an "X".

4. Words marked "rare," "obs.," "colloq." and "dial.," were indicated in the same way in the copied list.

The dictionary contained approximately 1,775 pages of words and definitions. Although several columns yielded no words because they were given over entirely to explanations of letters of the alphabet or to lengthy definitions of words in preceding columns, the completed list contained more than 5,000 words.

After a preliminary analysis of the list, all words marked "rare," "obs.," etc., and those which had been checked with an "X" were dropped immediately from further consideration. The sample was then checked against Thorndike's list, and all words common enough to be found in this list were eliminated. The revised list contained 1,535 words. These, together with the "upper five thousand" in Thorndike's list, constituted the sources from which the final reservoir of words was selected.

Evaluation of the words in these lists.- The two lists of words were submitted to six competent judges for their criticism. The judges were: two university professors of English who had expressed an interest in vocabulary tests, a professor of educational psychology, two experienced teachers of English in a senior high school, and a graduate student who had made an extensive study of vocabulary tests and was at the time engaged in constructing one. Each individual was asked to indicate: (1) which words, in his opinion, would be most appropriate for a vocabulary test designed for use with students ready to enter college, (2) which

words might be of value, and (3) which words seemed obviously unfit for such a test. The writer also made an independent check on the word lists.

A tally was then made of the number of times each word had been checked as "usable," "doubtful," and "undesirable." Words checked as usable by three or more judges were retained, except in a few instances where two or more of the remaining judges had checked them as undesirable. In that case the word was omitted in spite of the favorable judgments. By this method 986 words were selected from the 6,535 words which had been submitted to the judges. The writer then chose 14 words from Thorndike's list which had been checked favorably by two persons and "doubtful" by others, merely to round out the list to 1,000 words. These words, arranged alphabetically, became the reservoir from which the test words were drawn. It was formulated in accordance with the theory which Inglis proposed. It was intended to be a sample of those words which he so aptly described as "the intelligent reader's vocabulary", a list of words containing neither the easiest and most common words nor any which are technical or especially limited in their application.

Procedure followed in constructing trial forms.- Alternate words were taken from the reservoir of 1,000 words which has just been described. Of these 500 words, the 1st, 5th, ... 497th were set aside for Form A of the experimental forms. The 2nd, 6th, ... 498th words were used for Form B; the 3rd, 7th, ... 499th for Form C; and, the 4th, 8th, ... 500th for Form D. Thus four lists of 125 words each were constructed.

Test exercises were then designed in accordance with the general plans which have already been described; that is, each test word was used in a short phrase, clause, or sentence, and its synonym or definition was placed among four plausible-sounding but unrelated words or phrases.

Description of the trial forms.- The directions asked the student to select and underline the best definition for the test word, and to place its number in the parentheses at the right. He was asked to indicate his answer by underlining the word as well as by copying its number because in previous experimental work numbers were sometimes written illegibly and, in a surprisingly large number

of cases, students copied the number of the test exercises instead of the number of the answer to it. The directions for the test were quite brief, and occupied only a small portion of the first page. Since the test was to be practically untimed, test items occupied the remainder of the page. The following exercises from Form A illustrate the nature of the test:

16. A candid opinion

- (1) reasonable (2) frank (3) prejudiced (4) well-founded
(5) far-fetched ()

100. Spurious rumors.

- (1) many (2) dangerous (3) injurious (4) rapidly-spreading
(5) unauthentic ()

The general plan for the test was an outgrowth of the experimental work with the Inglis test which has already been described.

Usable test exercises could not be designed for a few of the words. It was either impossible to explain their meaning by a synonym or short phrase, or, plausible "catch words" could not be found to accompany them. Consequently the experimental forms did not have 125 items each, but instead there were 114 exercises in Forms B, C, and D and 115 exercises in Form A. It was hoped, however, that out of the 457 exercises in the four experimental form 200 exercises could be selected for the four 50-word forms of the final test.

Administration of the trial forms.- All four forms of the test were administered to three groups of students. In the first group there were 48 students in an extension class - for the most part college freshmen and sophomores. In the second group there were 49 college freshmen, and in the third 14 high-school seniors. A few students failed to finish all forms of the test. By discarding all unfinished papers 95 comparable sets remained. The reliability of the 114-word forms was approximately .90. The means were between 65 and 70 in each case, indicating a very satisfactory level of difficulty.

Evaluation of the test exercises.- A tally was first made of the number of times each of the 457 test items was missed and the percentage of difficulty for each item was computed. The second step in the procedure was the selection

of the test papers of the 25 best students (with respect to vocabulary) and the 25 poorest students. These groups were selected on the basis of total score on the four forms of the test, it being assumed that the total score on the 457 items was sufficiently accurate as a measure of vocabulary ability. A check was then made of the number of students in each criterion group who got the item right and wrong respectively. "V" was then computed for each word, using the formula described on page 18, and the entire list of words and their values were assembled in tabular form. Table XII illustrates the arrangement.

TABLE XII
DIFFICULTY AND VALIDITY OF THE FIRST FOUR ITEMS OF FORM A

Number of Exercise	Word	Upper 25 Per Cent		Lower 25 Per Cent		Per Cent of Difficulty	V
		Right	Wrong	Right	Wrong		
1	aback	11	14	12	13	.56	-.04
2	acute	24	1	15	10	.22	.36
3	advert	19	6	18	7	.24	.04
4	aground	21	4	14	11	.26	.28

The results of the tabulation were then transferred to a scatter diagram showing the number of words available at each level of difficulty and at each level of validity as in Diagram 1. A heavy line has been drawn through this diagram.

Per V Cent	-	0	0-9	10	20	30	40	50	60	70	80	Total
90	0	1	0	0	0	0	0	0	0	0	0	1
80	6	3	3	3	1	0	0	0	0	0	0	16
70	3	2	5	9	5	1	1	0	0	0	0	26
60	1	0	7	8	13	3	7	1	1	0	0	41
50	2	2	3	4	10	11	7	2	4	0	0	45
40	1	1	3	7	13	6	20	4	7	2	0	64
30	0	1	2	5	8	9	23	11	8	1	0	68
25-29	2	0	2	3	6	5	11	3	1	0	0	33
20-24	1	0	2	4	12	8	3	2	0	0	0	32
10	0	3	13	20	22	14	5	0	0	0	0	77
0	3	6	22	15	8	0	0	0	0	0	0	54
Total	19	19	62	78	98	57	77	23	21	3	0	457

Diagram 1. - Classification of 457 words with respect to their relative difficulty and validity

The words whose values lie above and to the right of the line are the more desirable ones. By the very nature of the validity formula words with very high or low "difficulty" values tend to have low "V" values. Hence, lower "V" values may be regarded as more satisfactory when the percentage of difficulty is high than when it is average or below average. The available stock of the better words as summarized from Diagram 1 may be noted from Table XIII.

Construction of 50-word forms.- In selecting words from each "difficulty group" an attempt was made to pair them with respect to validity. Words were shifted about freely from one form to another. A word from Form "A" in the original 114-word forms might be used in Form "D" of the 50-word forms. Cumulative totals were kept as the lists were being prepared, so that if one form showed a greater

TABLE XIII
SUPPLY OF USABLE WORDS FOUND IN THE ORIGINAL TEST FORMS

Per Cent of Difficulty	Number of Words with Satisfactory "V" Value	Approximate Number of Words Available for Each 50-word Form
80 or more	7	2
70-80	21	5
60-70	33	8
50-60	34	8
40-50	52	13
30-40	52	13
25-29	4	1

total than another at one difficulty level, the difference could be eliminated by a shift of words or by compensatory selection at the next level. The data in Table XIV show how the forms compared when completed.

TABLE XIV
COMPARISON OF FOUR FORMS OF A VOCABULARY TEST
WITH RESPECT TO DIFFICULTY AND VALIDITY

Form	Total of Difficulty Column (Decimals Omitted)	Total of "V" Column (Decimals Omitted)
A	2565	1832
B	2572	1831
C	2579	1828
D	2572	1828

After four 50-word forms had been thus arranged another form was constructed, using those words which fell just below or to the left of the line in Diagram 1. That is, using the exercises which seemed to be most nearly satisfactory - yet below the accepted standard - Form Z was constructed. The total of the "difficulty column" on this form was 2262 as compared with approximately 2570 on the others, and the total of the "V" column was only 1016 in this instance as compared with approximately 1830 on the others. Form Z was constructed in order that the best exercises in it might be used in the place of those which proved unsatisfactory in the regular test forms as the work of experimentation and revision was carried forward.

The general make-up of the revised tests was identical with that of the original 114-word forms.

Evaluation of the first 50-word forms.- The 50-word vocabulary tests were first administered in a class of 335 high-school seniors. However, it was not possible to have each student take two forms of the test. Form A was given to 45 students, Form B to 76, Form C to 87, Form D to 84, and Form Z to 43. The second group tested was a class of 33 high-school seniors. In this case, each student took all forms of the test on consecutive days.

In every case the medians were lower than had been anticipated. Since the criterion of 50 per cent difficulty had been used as a guide in constructing the tests, the medians were expected to fall at approximately 25, the mid-point of the possible range of scores. The medians presented in Table XV show how far that assumption was in error. To be sure, the percentage of difficulty was based almost

TABLE XV

MEDIAN SCORES OF TWO GROUPS OF HIGH-SCHOOL SENIORS ON
FIVE FORMS OF A 50-WORD VOCABULARY TEST

Group	Form				
	A	B	C	D	Z
1st group	9.9	13.6	10.8	12.4	17.7
2nd Group	15.8	17.7	16.0	15.7	21.6

entirely on the test papers of college freshmen and sophomores, whereas these scores were made by high-school seniors. The difference, however, was rather large to be explained entirely on that basis. It is doubtful if the groups originally tested would have made average scores of 25 on these forms. At any rate, they were too difficult for the students to whom they were administered.

The medians for the first group were somewhat lower than those for the second. This was probably due to the way the test was administered more than to a difference in the ability of the two groups. In the first instance the test was administered to the 335 students assembled in an auditorium. When it appeared that most of the students had finished, the work was stopped. As a matter of fact, many were not quite through working. Moreover the working environment in such a large group is not as good as that in a classroom with 33 students. Under similar conditions it is likely that the medians would have been approximately the same.

Two conclusions were drawn from the above medians: (1) that the tests were too difficult (the level of difficulty of Form Z was more appropriate), and (2) that Forms A, B, C, and D were approximately equal in difficulty. To be sure, the median on Form A with the first group was lower than any of the others, but it was used with only 45 students in that instance and chance factors could easily have been responsible for the observed difference. The medians were nearly equivalent when based on the scores of the same students. Each of the foregoing conclusions entered into the revision of the tests. The first implied that the tests should be made easier; the second, that if individual items in the various forms were paired with respect to difficulty, the completed test forms would probably be comparable.

1

Procedure used in revising the tests. - The percentage of difficulty and

1
A copy of the vocabulary test Form A is presented in the Appendix, page 189.

the validity of each of the 250 exercises were recomputed. Tables were prepared setting forth the original and later values in parallel columns. Four forms were then drawn up, using words from Form Z to replace unsatisfactory words in the remainder of the forms. Items were selected which showed the highest individual validity in both instances. The level of difficulty was equated as before and reduced to approximately the same general level as that of Form Z. The total of the "difficulty column" for each of the revised forms was as follows (decimals omitted): Form A, 2291; Form B, 2305; Form C, 2299; and Form D, 2297. The corresponding figure for Form Z in the original series was 2262, and for Forms A, B, C, and D, approximately 2570. Table XVI summarizes the number of words at each difficulty level in the new test forms and the corresponding number in the first 50-word forms. There were fewer difficult words in the revised forms and more words at the lower end of the scale.

TABLE XVI

NUMBER OF WORDS AT EACH LEVEL OF DIFFICULTY IN THE FIRST
50-WORD VOCABULARY TESTS AND IN THE REVISED FORMS

Per Cent of Difficulty *	Number of Words in First 50-Word Forms	Number of Words in Revised Forms
80 +	2	1
70 - 80	5	4
60 - 70	8	6
50 - 60	8	7
40 - 50	14	11
30 - 40	13	14
20 - 30	--	5
15 - 20	--	2
Total	50	50

*
The percentage of difficulty in each case is based on the original computations.

One additional set of facts will be presented largely as a matter of interest. All the test words were either nouns, adjectives, or verbs, and the various forms contained approximately the same number of each. In Form A there are 21 nouns, 22 adjectives, and 7 verbs; in Form B, 22 nouns, 20 adjectives,

and 8 verbs; in Form C, 24 nouns, 19 adjectives, and 7 verbs; and in Form D, 25 nouns, 18 adjectives, and 7 verbs. That the three parts of speech were so equally represented in the test forms was accidental, but perhaps there is merit in the arrangement.

Reliability of the revised test.- Forms A and B of the final vocabulary test were first administered to 27 college freshmen. The correlation between the two sets of scores was $.837 \pm .04$. Forms C and D were given to a class of 38 college freshmen and sophomores. The correlation in this instance was $.916 \pm .02$. The test was then administered to 191 college freshmen. Each student took all forms of the test. The reliability coefficients obtained from the scores of these students are presented in Table XVII. The results were consistent and the group was large enough to make them significant. These coefficients are somewhat

TABLE XVII
RELIABILITY COEFFICIENTS FOR FOUR FORMS OF
THE "TEST OF READING VOCABULARY"

Form	r $\pm$ P. E.			
	A	B	C	D
A	-----	$.823 \pm .016$	$.835 \pm .016$	$.821 \pm .016$
B	-----	-----	$.828 \pm .016$	$.827 \pm .016$
C	-----	-----	-----	$.834 \pm .016$
D	-----	-----	-----	-----
Number of cases		191	191	191

lower than any obtained with other groups. If they represent minimum reliability for the test - as they seem to do - they need not be deplored. Forms A and B were used later with 157 college freshmen and the reliability coefficient in that case was $.873 \pm .014$. The foregoing coefficients, ranging between .82 and .92, indicate that the final forms of the test are moderately reliable.

Equality of the forms.- The relative difficulty of the four forms of the test is indicated by the data in Table XVIII. These data were derived from the scores of the 191 college freshmen. The means and standard deviations were practically identical, showing that the test forms were comparable. The comparatively

high reliability of the test adds to the significance of this comparison.

TABLE XVIII

COMPARISON OF THE MEAN AND STANDARD DEVIATION FOR
FOUR FORMS OF THE "TEST OF READING VOCABULARY"

	Form			
	A	B	C	D
Mean	18.6	18.6	19.1	20.3
Standard Deviation	8.3	7.8	7.8	8.0

Other evidences of the value of the test.- Additional facts will be presented in later sections of this report confirming and supplementing those presented in the foregoing paragraphs.

Development of a Comprehension Test

What is comprehension?- The first problem in constructing a comprehension test was that of discovering a tenable working definition for the word "comprehension". A review of the definitions given by psychologists and students of reading problems and an analysis of the activities demanded by comprehension tests already constructed seemed to indicate that comprehension could be measured in terms of four abilities: (1) ability to determine total meaning, (2) ability to determine correct relationships between ideas, (3) ability to answer fact questions, and (4) ability to follow directions. This working classification of the measurable activities indicating the quality of comprehension emphasizes active mental processes.

It was not assumed that the most perfect measure of these activities would constitute a complete measure of "comprehension". This analysis merely suggested an empirical definition of a complex ability; it outlined the scope and limitations of the proposed comprehension test.

How may these abilities be measured?- The ability to read and follow directions was believed to be so essential for satisfactory performance on the test battery as a whole that a separate test of this ability was deemed unnecessary. The check on comprehension used in connection with the rate test was

expected to indicate to some extent the student's ability to grasp the total meaning of a selection, but its value was as yet undetermined. Therefore, the need seemed to be for a series of exercises that would measure the reader's ability (1) to interpret TOTAL MEANING, (2) to grasp correct relationships - ORGANIZATION, and (3) to answer FACT QUESTIONS. In view of the facts stated above, "total meaning" appeared to require less emphasis than either of the other factors.

General plan for the test.- To measure the three more or less distinct processes which have been designated by the titles "total meaning", "organization", and "fact questions", either of three courses were open: (1) the test could be divided into sections, each devoted to the measurement of one phase of comprehension, (2) each paragraph in the test could be followed by test exercises of the three types, identical in number and arrangement in each instance, or (3) test exercises of each type could be inserted into each test form with no particular attention to placement. The second plan was adopted. The test would not be a real aid in diagnosis unless the exercises dealing with each separable ability were placed in such a way that they could be identified quickly. The second plan would place them so. At the same time, this plan seemed to require a smaller quantity of test material in proportion to a given number of test exercises than a three-part test in which all the exercises in each part were of a single type. From the standpoint of economy of time in administering the test, the latter consideration was an important one.

Technique selected for the test.- The multiple-choice technique was selected as an appropriate one for measuring the reader's ability to understand¹ the total meaning of a paragraph. Multiple-choice exercises can be used effec-

1

For evaluations of this technique see: W. Asker, "Reliability of Tests Requiring Alternative Responses," Journal of Educational Research, IX (March, 1924), 234-40; W. H. Burton, "A Contribution to the Technique of Constructing 'Best-Answer' Tests," Elementary School Journal, XXII (June, 1925), 762-70; H. H. Hahn, "A Criticism of Tests Requiring Alternative Responses," Journal of Educational Research, VI (October, 1922), 236-40; and H. H. Remmers, et al., "An Experimental Study of the Relative Difficulty of True-False, Multiple-Choice, and Incomplete Sentence Types of Examination Questions," Journal of Educational Psychology, XIV (September, 1923), 366-72.

tively when judgment and discrimination are important elements in the response. The plan - also used in other tests - of having the student select from a series the phrase or sentence which best expresses the central thought of the paragraph seemed appropriate in this instance.

The second type of measurement anticipated was one dealing with the reader's ability to organize ideas - to select important ideas and see them in relation to one another. None of the devices used by previous test makers seemed to merit imitation. After studying a number of the organization exercises in other tests, a new technique was developed, combining some of the features of other devices but differing from them all. It may be described briefly as a modified form of the multiple-choice technique. The plan was to have the student select from a series of short sentences or phrases a few related ones. He would then arrange these related ideas in correct order to show how they were related. He would indicate this arrangement by placing the numbers of the sentences selected in a schematic outline provided. The number of sentences or phrases in the series, the number of related ideas to be used, and the method of scoring were determined by trying out different arrangements. The device as ultimately used in the test will be described and illustrated later.

In order to measure the third phase of comprehension, the ability to answer fact questions, a modified form of the true-false technique was developed.¹ The ordinary form of the true-false exercise was not adopted, because with a small number of such exercises, chance errors due to guessing may invalidate the results. Moreover, it seemed reasonable - particularly in the case of college students - to ask the reader to judge each statement in terms of three possibilities, rather

I

For evaluations of the true-false technique see: M. F. Fritz, "Guessing in a True-False Test," Journal of Educational Psychology, XVIII (November, 1927), 558-61; H. C. Lehman, "Does it Pay to Change Initial Decisions in a True-False Test?" School and Society, XXVIII (October, 1928), 456-58; H. H. Remmers, et al., op. cit.; H. M. Roberts, and G. M. Ruch, "The Negative Suggestion Effect of True-False Tests," Journal of Educational Research, XVIII (September, 1928), 112-16; and C. C. Weidemann, "Limitations of the True-False Statement," Journal of Educational Method, VII (February, 1928), 214-15.

than two; to decide whether a given statement was true, false, or unrelated to the paragraph. Accordingly this modification was introduced. The number of fact-question exercises to be used was determined by experience with trial forms of the test.

Selection of test paragraphs.- Twenty-seven paragraphs were assembled as the initial step in the construction of comprehension test forms. They were drawn from standard books and magazines. Most of the passages selected were drawn from continuous discourse and for that reason seemed fragmentary when used alone without being modified. Therefore, adaptations were made freely in them to produce complete, unified paragraphs.

The material was selected on the basis of maturity of thought, simplicity of style, interest, and originality in point of view. The paragraphs differed greatly with respect to content, sentence structure, type of discourse, and vocabulary. They varied in length from about 300 to 400 words.

Trial use of single paragraphs.- Test exercises of the three kinds described were prepared for about half the paragraphs originally selected. Each paragraph was handled as a separate unit. The number of exercises constructed differed from paragraph to paragraph, especially those pertaining to organization and fact questions.

A number of students were then asked to do one or more of the paragraphs, working in an individual test situation. High-school seniors, college students, and graduate students read the paragraphs, did the exercises, and criticized both the materials and the exercises. Each paragraph was used with ten or more individuals. Many defects were corrected on the basis of their individual reactions. Entire paragraphs and many test exercises were discarded because of weaknesses discovered in this manner.

On the basis of this preliminary work, a number of important decisions were reached as to the nature of the comprehension test as a whole. First of all, the form for the organization exercises was developed. Exercises using three related ideas in a series of five or six sentences or phrases proved most satis-

factory, for more than three ideas which could be arranged in a uniform type of outline were very difficult to formulate for certain paragraphs. Various methods of scoring these exercises were also considered. The simplest method, namely, the point for point unweighted score, seemed as accurate and valuable as any system of weighting that was tried. Hence to measure organization, a three-point exercise was developed in which the student selected three related ideas out of six that were listed and arranged them to show how they were related.

Various forms of response to the fact-question exercises were tried out, as the single paragraphs were used with individuals. The simplest method found was that employing the three symbols: plus, minus, and zero (+, -, and 0). Students found the symbols easy to use and they were much easier to score than either words or letters. Trial and observation also led to the adoption of six fact-question exercises as the number to be used with each paragraph. Six exercises of the true-false type, even in this modified form, seemed few enough at best. More than six exercises would, in many cases, have made it necessary to use statements too obviously true, false, or unrelated.

Each student's working time on the various paragraphs was recorded. The average working time on one paragraph was about five minutes. Therefore, plans were laid for trial forms with five test paragraphs in each. Using five paragraphs a working time of thirty minutes - the amount planned for this part of the test battery - would be a fairly generous time allowance. The question was, Would a test of that length be sufficiently reliable?

All the foregoing decisions were based on subjective judgments. That is, no attempt was made to demonstrate by mathematical formula why ten exercises should be used with each paragraph, or why unweighted point-scores should be employed. The logical advantages of a five-paragraph test have been given. Likewise, the principal considerations involved in developing the various types of comprehension exercises have been presented but the correctness of these decisions cannot otherwise be demonstrated.

The three phases of comprehension are not equally represented in the test score, since with each paragraph there is one total-meaning exercise, three organization exercises, and six fact-question exercises. No particular theory or defense of that division of emphasis will be offered. It is not implied that the ability to organize ideas is three times as important as the ability to understand the total meaning of a paragraph; nor, that the ability to answer fact questions is twice as important as the ability to organize ideas. Each device was considered in terms of how well it seemed to cover the content of the paragraph rather than in terms of the number of points that it contributed to the total score. As a matter of fact, the various abilities are not so unequally represented if total meaning is also measured by the comprehension part of the rate test, as it seems to be. Moreover, if the test allowed exactly the same number of points for each of the three phases of comprehension involved, there would be no assurance that a better measure of comprehension was thereby secured.

Some of the questions discussed in this section of the report could be studied advantageously in elaborate and carefully controlled experiments. In this instance it was impossible to do so if other phases of the investigation were to be carried out as planned. Since the latter seemed the more important in terms of the study as a whole, the empirical methods described above were substituted for the more thorough methods which otherwise would have been warranted.

Construction of two trial forms. - Seven of the paragraphs used in the work just described were selected for use in the first test forms. These contained several distinct types of paragraphs. Two pairs of them, however, were very similar. One pair dealt with a social science topic, treating it from two different points of view. The other pair discussed two simple philosophical questions. These four paragraphs were selected as paragraphs four and five for the two trial test forms that were being assembled. The three other paragraphs selected to complete one of the trial test forms were quite dissimilar. It seemed logical that the paragraphs in the various test forms should parallel each other,

in so far as possible, with respect to general type of discourse, nature of thought content, plan of organization, and general level of difficulty. Therefore, three other paragraphs were prepared to resemble as closely as possible the corresponding paragraphs already selected for the first trial form. Test exercises were prepared for them, and like the earlier paragraphs, they were tried out first with individuals.

The ten paragraphs were then arranged in parallel series, mimeographed, and assembled for trial use with groups of students.

¹
Description of the test forms. -- The test was entitled A Test of Accurate Interpretation in Reading to distinguish it from the comprehension part of the rate test. The directions for the test were rather long and perhaps somewhat bewildering in spite of every effort to make them simple. Aside from the fact that they were rather hard to understand, the directions were justly criticized on the ground that they were time-consuming. The only defense offered for a technique requiring such elaborate directions is that the time lost with the directions is justified by economy of time in the administration and scoring of the test.

Since the nature of the comprehension test has been explained at some length in previous paragraphs, it can be described best at this point by giving an illustration. The following paragraph and exercises are taken from the page of directions. The answers to the exercises have been indicated to show how the student's responses are given.

SAMPLE PARAGRAPH. -- Advertising has changed remarkably in recent years. Its methods used to be quite simple and crude as compared with the highly organized departments and carefully planned advertising campaigns which feature modern business. For example, I remember the leading merchant in a little town out West. His store stood on the town's busiest corner, where his three sons helped him sell drygoods. Every Saturday noon in winter he used to get up on the roof of his store and throw a new overcoat into the street. The crowd scrambled for it. That was advertising in those days. Incidentally, one of his sons was supposed to be quick enough to grab the coat and rush it back to the shelves to be thrown down again next Saturday. It was a bad day for the boys if they let an outsider get away with that decoy coat!

¹

The test in its original form was so nearly identical to the final forms that only a copy of the latter is presented in the Appendix.

- I. The western merchant is portrayed as (1) opposed to advertising
(2) a miserly money grabber (3) a typical advertiser of his day
(4) a man with modern business ideas (5) a harsh father (3)

II. Make a partial outline using three of the following phrases:

- | | |
|--|----------------|
| (1) Saturday noons) | I. (5) |
| (2) Grab the coat.) | 1. (6) |
| (3) <u>A decoy coat</u>) | a. (3) |
| (4) Back to the shelves.) | |
| (5) <u>Advertising methods</u>) | |
| (6) <u>Former methods simple and crude</u>) | |

III. Indicate whether these statements are true, false, or unrelated:

- | | |
|--|-----|
| 1. Advertising methods have been changed considerably | (+) |
| 2. Modern advertising is highly organized | (+) |
| 3. The western merchant used modern advertising methods | (-) |
| 4. The merchant gave away a coat each Saturday in the year. | (-) |
| 5. Giving away merchandise is not good advertising. | (0) |
| 6. When the coat was thrown into the street one of the merchant's sons
tried to recover it. | (+) |

The foregoing illustration shows that the test exercises were entirely objective. The student's method of response and the method of scoring were quite similar to those employed in both the rate and vocabulary tests. The student was required to do practically no writing. All answers were to be indicated by placing numbers or symbols in parentheses provided for that purpose. The score on the test was the number of correct responses, ranging in each case from 0 to 50. In order to distinguish it from certain other scores relating to comprehension, the score on the Test of Accurate Interpretation in Reading will be referred to, from this point on, as score "C_g".

The test paragraphs were written in relatively simple style, but in language appropriate for adult readers. No technical words or expressions appeared in them. They were not scaled according to difficulty but, as has been pointed out, each paragraph in one form was paired quite definitely with a paragraph in the other form. Each paragraph in a given form differed from all the others with respect to style or type of discourse. In each form the first paragraph was a simple narrative. The second was a straightforward argument intending to show the fallacy of a widely accepted belief. The third was a vivid description. The fourth paragraph was a discussion of the topic "man's social

heritage" such as might be encountered in a social science textbook. The last was a discussion of a concept which was abstract and more or less philosophical in tone.

The forms were designated 100 and 200. Each paragraph was numbered 101, 102 105, and 201, 202 205. Since each paragraph was placed on a separate page and numbered in this fashion, the interchange of paragraphs and the evaluation of separate paragraphs were facilitated.

Relative difficulty of Forms 100 and 200.- These test forms were first administered to 51 seniors in a suburban high school. These students were selected from the upper third of a class of some 350 students. It is, therefore, reasonable to assume that the group was somewhat more homogeneous with respect to comprehension in reading than an unselected group would have been. The tests were also given to 32 college freshmen and sophomores. With the high-school group the means were 33.2 and 34.5, respectively. With the college group the means were 35.0 and 33.3. With respect to general level of difficulty, better results could scarcely have been expected. The means were close, and they fell at a very satisfactory point in the 50-point scale.

Reliability of the trial forms.- The reliability of the test with the college students was only .614 and with the high-school group only .604. Such a result was not particularly gratifying, but the fact that the high-school group was quite homogeneous would tend to lower the correlation in that instance. Moreover, experimentation with the rate and vocabulary tests had indicated that the validation of individual test items might result in a significant increase in the reliability of the test as a whole. Therefore, attention was directed toward that project. The validity of each exercise was computed by the use of the formula¹ already described. Many individual test items were discovered with zero or negative "V" values. Practically all other items showed low positive values. Each item which showed a negative value, zero value, or very low positive value was examined carefully in an effort to discover why it failed to contribute to

¹

See p. 18.

the test. A number of items were too easy; others were too hard; others were ambiguous in meaning.

Revision of Forms 100 and 200.-- On the basis of the foregoing analysis, the trial forms were revised. The test paragraphs were kept without any major changes. The directions were shortened by about one-fourth and were clarified at certain points.

The most important revisions, however, were those of the test exercises. Two faults were so frequently encountered that they will be pointed out as a warning to those who may wish to construct exercises of this type. The first was, that more than one valid outline was sometimes possible. In the organization exercises there were supposed to be three related ideas and three ideas unrelated to any others. Students sometimes discovered relationships - perfectly logical ones - which had not been suggested in any of the work with single paragraphs. Because of such new interpretations, two or more valid outlines could sometimes be made from the ideas listed. The second fault was, in a way, similar to the one mentioned above. It had to do with the supposedly unrelated, true-false items. Students' responses sometimes showed clearly that they saw an implied relation between the statement and the content of the paragraph although none was intended. The low validity of single test items led to the discovery of faults such as these.

When the tests were again tried, they seemed satisfactory for the purposes of this study. Had it been possible to do so without omitting other necessary phases of the investigation, the comprehension test could doubtless have been improved by further study and refinement. However, the means were again very close and the reliability was considerably higher. Forms 100 and 200, although admittedly incomplete, were regarded as useful measuring instruments. Data relating to their value will be given presently.

The construction of Forms 300 and 400.-- Four of the test paragraphs used in the earlier work with single paragraphs were used in Form 300 and 400; two in the field of social science, and two in philosophy. Paragraphs 301, 302, 303, and 401, 402, and 403 were written to parallel, as closely as possible, the correspond-

ing paragraphs in Forms 100 and 200. Test exercises were prepared and the paragraphs were tried with five or six individuals. The tests were then revised, assembled for use with groups, administered to one class of high-school seniors, and the exercises revised once more. The comprehension test was the last of the three to be developed and when it reached this stage of completion, it was necessary either to use the test in this form or to discard it entirely. No more time could be devoted to its further refinement if other phases of the study were carried out. Therefore, the four forms were subjected to a limited amount of comparative use to determine whether they should be used in the later phases of the study or whether this test should be laid aside. Tentative statements as to the validity of each item of the test, the reliability of the various parts and of the test as a whole, and the relative difficulty of the various forms are presented in the following paragraphs.

Validity of the comprehension test.- There is less objective evidence of the validity of the comprehension test than of either the rate or vocabulary tests. Its merits and limitations must be judged largely by the soundness of the assumptions which have been presented. Of course that is true of any test, to some extent. No matter what statistical methods of validation have been employed, a test must be judged in no small measure by the assumptions on which it rests. The facts presented in the following paragraphs with reference to the validity of individual test items and the reliability of the test supplement this subjective evaluation. Moreover, data presented in the following chapter show that the test does not correlate closely with intelligence nor with other reading abilities. Nevertheless, none of these facts nor all of them combined give conclusive proof that the test is a valid measure of comprehension. It is valid only to the extent that its objectives are sound - to the extent that it measures accurately three factors which taken together give adequate insight into the quality of the reader's comprehension.

Validity of single items.- Each exercise on the comprehension test was evaluated separately using the formula described on page 18 with thirty students in each criterion group. The indices of validity for the four forms of the test appear

in Table XIX. The table is to be read as follows: The first test exercise in the first paragraph of Form 100 showed a validity (in terms of "vw") of .20; the second

TABLE XIX

VALIDITY OF THE ITEMS OF THE "TEST OF ACCURATE INTERPRETATION IN READING"

Paragraph Number	Item Number									
	1	2	3	4	5	6	7	8	9	10
101	.20	.13	.20	.33	.00	.20	.13	.27	.07	.13
102	.13	.27	.20	.33	.13	.13	-.07	.33	.20	.07
103	.07	.27	.13	.07	.13	.00	.20	.20	.13	.47
104	.33	.13	.40	.07	.07	.27	.13	.13	.20	.27
105	.07	.33	.47	.20	.27	-.07	-.07	.13	.13	.27
201	.13	.00	.60	.00	.00	.47	.33	-.20	.07	.27
202	.13	.20	.33	.33	.07	.13	.13	.13	.00	.00
203	.33	.07	.33	.20	.20	.00	.07	.07	.00	.07
204	.07	.13	.20	.20	.07	-.07	.13	.07	.47	.07
205	.00	.20	.13	.47	.13	.40	.33	.20	.13	-.07
301	.27	.47	.27	.27	.27	.00	.13	.00	.07	.13
302	.20	.13	.13	.20	.00	.20	.33	.00	.00	.13
303	.13	.40	.60	.47	.07	.40	.20	.13	.00	.07
304	.00	.33	.47	.27	.47	.40	.00	.33	.27	.07
305	.13	.07	.20	.47	.13	.47	.40	.47	.40	.13
401	.33	.13	.27	.47	.20	.00	.07	.13	.20	.27
402	.27	.27	.33	.33	.07	.13	-.07	.00	.13	.20
403	.13	.47	.20	.27	.13	.20	.13	.00	.07	.07
404	.20	.60	.27	.13	.00	.27	.00	.40	.13	.20
405	.13	.33	.47	.13	.13	-.07	.20	.27	.13	.07

item in the same paragraph, .13; and so on. Many of the values were low, especially those for the true-false exercises (items 5 to 10 inclusive). Eight of the 200 exercises had negative values; 22 items had zero value. In practically all cases the values were lower than .50.

The low validity of the items on this test was discovered early in the course of its construction. In revising the test, the validity of single items was taken into account and every effort made to improve the poorer ones. Nevertheless, mere positive values seemed to be the only attainable objective. There was a much larger percentage of zero and negative values for the items of the original test forms, indicating that the revisions were not in vain. Moreover, in each test form there are a goodly number of items with satisfactory "vw"

values. Further revision and refinement of the items would probably increase their individual validity to some extent, but the possibility of increasing it to any great extent seems remote. Apparently, test exercises of certain types tend to have lower validity than those of other types.

Reliability of the test.— The reliability of the Test of Accurate Interpretation in Reading will be considered from two angles: (1) the reliability of the various parts of the test, and (2) the reliability of the test as a whole.

In order to determine the reliability of the various parts of the test — total meaning, organization, and fact questions — the test papers of 100 college freshmen were studied in detail. Forms 200 and 300 were involved. The students' scores on each part of the test were determined and the various part-scores were correlated. The resulting coefficients appear in Table XX. The least reliable of the three parts was that dealing with total meaning. However, the possible range

TABLE XX

RELIABILITY OF VARIOUS PARTS OF THE "TEST OF
ACCURATE INTERPRETATION IN READING"

Parts Correlated	Number of Cases	$r \pm P. E.$
Total Meaning, Form 200 with Total Meaning, Form 300	100	.276 $\pm$.062
Organization, Form 200 with Organization, Form 300	100	.386 $\pm$.057
Fact Questions, Form 200 with Fact Questions, Form 300	100	.595 $\pm$.043
Paragraph 201 with Paragraph 301	100	.242 $\pm$.064
Paragraph 202 with Paragraph 302	100	.207 $\pm$.065
Paragraph 203 with Paragraph 303	100	.231 $\pm$.064
Paragraph 204 with Paragraph 304	100	.369 $\pm$.058
Paragraph 205 with Paragraph 305	100	.374 $\pm$.058
Paragraph 201 with Paragraph 202	100	.304 $\pm$.061
Paragraph 202 with Paragraph 203	100	.193 $\pm$.065
Paragraph 203 with Paragraph 204	100	.232 $\pm$.064
Paragraph 204 with Paragraph 205	100	.302 $\pm$.061

in score on this part of the test is only 0 to 5. Therefore, the product-moment method of correlation is only a crude measure of relationship in this instance.

The coefficient listed is positive and more than four times its probable error. With so few exercises, little more could reasonably be expected. The coefficient of contingency for the same set of scores was ¹.463. Although this value is not identical in meaning with a product-moment coefficient, it indicates a degree of reliability for the total meaning part of the test. The most reliable part of the test was that composed of fact questions - strangely enough, the part in which individual test items showed the lowest validity. This apparent inconsistency was probably due to the difference in the length of the various parts. Note that the reliability coefficients increase (.276 to .386 to .595) as the range increases (5 to 15 to 30 items).

Scores on a given paragraph did not correlate highly with those on the corresponding paragraph in another test form. The coefficients ranged from .207 to .374. Again this low correlation is not surprising for each test paragraph is followed by only 10 items. The second paragraph was least reliable. The last two paragraphs were considerably more reliable than either of the first three. The various paragraphs in Form 200 showed about the same degree of correlation with each other as that of corresponding paragraphs in the different test forms, about .20 to .30.

The forms of the test were administered to four groups of students, using a different combination of forms in each case. Although the correlations based on these test scores warrant no sweeping generalizations, they do suggest the general level of reliability for this test. The coefficients are summarized in Table XXI. Three of them were higher than .70 and the fourth was only a little lower than .70.

Equivalent forms.- The extent to which the various forms of the test were comparable was also demonstrated, at least in part, by the same scores on

¹
For an explanation of the formula for the coefficient of contingency, $c = \frac{\sqrt{S-1}}{S}$, see Holzinger, op. cit., pp. 273-78.

TABLE XXI

RELIABILITY COEFFICIENTS FOR THE "TEST OF ACCURATE
INTERPRETATION IN READING", SCORE "Cs"

Forms Administered	Number of Cases	$r \pm P. E.$
100 and 200	60	.819 $\pm$.03
200 and 400	29	.736 $\pm$.056
100 and 300	37	.766 $\pm$.045
200 and 300	157	.665 $\pm$.03

which the foregoing correlations were based. The means for the various distributions are summarized in Table XXII. They were quite close in every case. From the standpoint of relative difficulty the test seemed quite satisfactory.

TABLE XXII

MEANS FOR FOUR FORMS OF THE "TEST OF ACCURATE
INTERPRETATION IN READING"

Number of cases	Mean on Form				Difference
	100	200	300	400	
60	28.4	28.2	----	----	.2
37	33.5	----	34.0	----	.5
29	----	32.9	----	33.5	.6
157	----	29.4	29.0	----	.4

Value of the comprehension test.— Although the data given in the foregoing paragraphs do not constitute a thorough evaluation of the comprehension test, they provide ample evidence of its usefulness. Constructed according to a plan that seems fundamentally sound, with test items showing positive validity in nearly every case, with reliability of about .70, and with forms approximately equal in difficulty, the value of the test is apparent. On the other hand, the data suggest the possibility and the desirability of further refinements, by means of which its value could doubtless be increased. The test was used in later phases of this study, its merits and its limitations being fully realized.

The Completed Battery of Silent Reading Tests

In the preceding sections of this chapter, the development of the Test

of Achievement in Silent Reading, the Test of Reading Vocabulary, and the Test of Accurate Interpretation in Reading have been discussed as though each was a separate test. Although, in one way, each test is complete within itself, in a larger sense, the three are quite interdependent. Therefore, the final section of the chapter deals with the tests as component parts of a larger unit, showing the place of each in the series, and evaluating the test battery as a whole. Omitting further description, this discussion begins with directions for administering and scoring¹ this test.

Administering the test.— One form of the rate test, the vocabulary test, and the comprehension test should be assembled into a single test booklet, placing them in the order named. At the beginning of the period the examiner should see that the test environment is as favorable as possible. The papers may then be distributed with the comment, "This is a test to see how well you can read. I shall pass to each of you a test booklet. As soon as you receive it, fill the blanks on the first page and wait for further directions. Do not open the booklet until you are told to do so."

Allow a reasonable length of time for filling the blanks. Then say, "The test has three parts, (1) a Test of Achievement in Silent Reading, (2) a Test of Reading Vocabulary, and (3) a Test of Accurate Interpretation in Reading. When you have finished one part of the test, close the booklet and wait quietly until the others have finished so that all may start on the next part at the same time."

The examiner should then read orally the directions for the Test of Achievement in Silent Reading appearing on the first page of the test blank and set the group to work as indicated there. He should note the time to the exact second when the students turn to page 3. After 100 seconds and again after 200 seconds the examiner will say, "mark; and keep right on reading". If possible, a stop watch should be used. If an ordinary watch is used it is of the utmost

¹

See pp. 6, 26, and 39; also a copy of the test given in the Appendix, pp. 185-97.

importance that the examiner note the starting time to the exact second. A satisfactory method for doing so is described in the following instructions. Hold the watch and a fountain pen ready while reading the material on page 2 and as soon as the students turn the page put a dot of ink on the crystal of the watch directly above both the second hand and the minute hand. Then determine and mark the position of the hands at the end of 100 seconds and 200 seconds. Since the position is different in each case, two dots may be used to indicate the position of the hands at the end of 100 seconds and three dots the position after 200 seconds. The examiner can then say, "Mark," just as the hands pass the points indicated. The ink spots show up clearly against a white background and they can be removed easily from the crystal of a watch.

Fifteen minutes should be ample time for the student to fill the blanks, master the directions, read the essay, and answer the questions on page 2. At the end of that time (or sooner if all have finished) the students may be asked to turn to the Test of Reading Vocabulary.

The directions should be read orally and the group set to work as indicated. Fifteen minutes should be allowed for the vocabulary test.

The students should then turn to the Test of Accurate Interpretation in Reading and follow through those directions carefully as the examiner reads them orally. It is important that the student should actually do the things which he is asked to do in the practice exercises. The examiner should read the directions well, emphasizing by tone of voice the important elements. Every student should understand what is wanted. The directions should not be given rapidly. It is well for the examiner to use about five minutes in going through them with the group. The final part of the test is untimed, but the student will seldom accomplish much after thirty minutes. Time may be called sooner if most of the students have finished the test.

Scoring the test.- The scoring key for the rate test has two essential parts. First of all, the sums obtained by making a cumulative line by line word

count are recorded in the left-hand margin of a regular test blank. These numbers are divided by two and the quotients inserted in the right-hand margin. This series constitutes the second and most important part of the scoring key. A student's 100-second rate score is found by locating the line in which his first mark appears and taking from the left-hand margin of the scoring key the number for that line. His 200-second rate score is found by locating his second mark and taking the number for that line as it appears in the right-hand margin. These are recorded on the student's test blank as Scores A and B, respectively. Only the 200-second rate score need be recorded unless the student reads so rapidly that he finishes before the end of the 200-second period. That is, if the student has marked twice, only the second mark need be considered and the score is obtained from the right-hand margin. If only one mark appears, the scorer should check with the student to find out whether he finished reading before the end of the 200-second period or whether he failed to hear the examiner say, "Mark," in one instance.

The comprehension part of the Test of Achievement in Silent Reading, the Test of Reading Vocabulary, and the Test of Accurate Interpretation in Reading are scored by comparing the numbers or symbols which the student has placed in the parentheses with the scoring key - a test blank with the exercises correctly answered. The scorer should watch for two common errors. Frequently the student underscores the correct answer but forgets to put the number in the parentheses. Again the student may underline the correct response but copy the number of the test item instead of the number of the response. An impossible answer such as 6, 8, or 14 is almost certain to indicate the latter type of error. In all cases of doubt the scorer should rely on the response which is underscored.

When the process of scoring reaches this point, four scores will have been determined: a rate score, a vocabulary score, and two comprehension scores. The latter have been designated as score "Cr" (score on the comprehension exercises of the rate test) and score "Cs" (the comprehension score on the separate comprehension test). Two steps remain: (1) the combination of scores "Cr" and

"Cs" into a single comprehension score, and (2) the combination of all the scores into a composite score. A method for deriving each of these scores is described below, together with an explanation of why that method was adopted.

Total comprehension score.-- As a matter of convenience, the total comprehension score will be called, score " C_t ". It is a derived score based on scores " C_r " and " C_s ". The simplest method of combining scores " C_r " and " C_s " is to add the two two point scores without weighting either of them. This method was tried but discarded in favor of one which proved to be more reliable, namely, that of doubling the score " C_r " and combining it with the score " C_s ". That the reliability of the score " C_t " was increased by giving double weight to the score " C_r " is not surprising, since the individual items accompanying the rate test have greater validity than those of the separate comprehension test. The difference in the methods is illustrated by the following case. The score " C_t " was computed for 157 students on two forms of the test using the two methods: (1) " C_r " plus " C_s ", and (2) twice the score " C_r " plus " C_s ". The reliability of " C_t " with " C_r " unweighted was $.722 \pm .026$. When " C_r " was doubled, the reliability of " C_t " was increased to $.771 \pm .022$.

Composite score.-- The derivation of the composite score described in this paragraph involves no attempt to weight accurately the rate, vocabulary, and comprehension parts which enter into it other than preventing any one part from overshadowing all the others. The composite score is obtained as follows: (1) the rate score is multiplied by ten, that is, the decimal point in the words-per-second score is moved one place to the right; (2) this result is added to the vocabulary score; and (3) to their sum is added the score " C_t ". The procedure will be illustrated presently. For the average reader the three parts of the test contribute to the composite score somewhat as follows: rate 40 points, vocabulary 30 points, and comprehension 48 points. Although they do not contribute equally, no part has disproportionate weight.

Tabulating the results.-- The most convenient method for recording the

scores seems to be the following tabular arrangement:

Name of the Student.	Comprehension			Rate	Vocabulary	Composite Score
	C _R	C _S	C _t			
Adams, B. E.	8	35	51	4.00	24	115
Brown, J. F.	9	40	58	4.24	30	130

This illustration also shows how the score "C_t" and the composite score are derived from the four basic scores. In the case of B. E. Adams, twice the score "C_R" plus the score "C_S" (16+35) is 51. Then, 51+40+24 = 115. Similarly for the next student: 18+40=58, and 58+42+30=130.

Interpretation of the test scores.— The rate score indicates the number of words per second which the student read. It should always be considered with reference to the accompanying "C_R" score. If the latter is unsatisfactory, so is the former. A rate score of 3.00 with a "C_R" score of 10 is better than a rate score of 6.75 with a "C_R" score of 3. In fact the examiner should be on the alert for rapid superficial workers who have looked "at the window but not through it" — who have seen words but not the thought.

The score "C_R" is but a brief and partial index of how well the student understood the meaning of the essay at the rate indicated by his rate score. It should be high. It is not a scale value but a test of mastery. A score of 5, or less than 5, is not satisfactory. With reasonable understanding of the essay the student should make a score of 7 or more.

The comprehension score "C_S" is a point score giving a partial measure of the student's ability to interpret sentences, questions, and paragraph meaning. An analysis of the test paper of a student will often reveal that he is weak in a particular phase of paragraph interpretation.

The score "C_t" is merely a better composite picture of the student's ability with respect to comprehension in reading than either "C_R" or "C_S" taken singly.

The vocabulary score helps to explain why difficulties exist in other instances. A low rate score may be due to the fact that the student does not recognize words readily. Difficulties in comprehension may be traced to the application of inaccurate word meanings. If, on the other hand, the reading vocabulary is well developed, the instructor may look immediately for other causes of difficulty.

The composite score may be used in broad group comparisons and in any placement groupings in which precision is unnecessary. It is useful for administrative purposes but of limited value for diagnostic or instructional purposes.

Evaluation of the Test Battery

Important criteria.- A good test must satisfy seven requirements: (1) it must measure a significant ability, (2) it must be valid, (3) it must be reliable, (4) comparable forms must be available, (5) it must be objective, (6) it must have dependable standards, and (7) its practical utility must be demonstrated. How does this battery of tests measure up according to these standards?

Significance of the ability measured.- The significance of reading ability was discussed at some length in Chapter I and a number of investigations were cited which bear witness to the importance of reading ability. Likewise, the significance of reading rate, of the ability to recognize word meanings, and of the power to interpret sentences and paragraphs readily and accurately has been discussed in connection with the development of the test forms. Neither the importance of reading ability as a whole nor of the three aspects of reading measured by these tests are seriously questioned by anyone familiar with student problems. Few would deny that this battery of tests satisfies the first of the seven criteria.

Validity.- Since the methods used in validating the various parts of the test have been discussed in previous sections of this chapter, the quality of the test battery in this particular needs no further comment.

Reliability.- As a final step in evaluating the reading test as a whole, two forms of the entire battery were administered to 664 college freshmen. The tests used were Forms 1 and 4 of the Test of Achievement in Silent Reading, Forms

A and B of the Test of Reading Vocabulary, and Forms 200 and 300 of the Test of Accurate Interpretation in Reading. The order in which the forms were given to half the class was reversed with the remainder of the students in order to eliminate practice effect. The data from this testing program are more extensive than those obtained in any other single program in this investigation. The study was limited to two forms of the test, but since the other forms are approximately equivalent to those used in this program, the results should be essentially correct as a description of all forms of the test.

The reliability of the test as a whole and of its various parts is shown in Table XXIII. The reliability of the composite score was almost .90. The various

TABLE XXIII
RELIABILITY OF A BATTERY OF SILENT READING TESTS

Score	$r \pm P. E.$
Composite Score	.892 $\pm$.005
Rate	.747 $\pm$.011
Vocabulary	.908 $\pm$.005
Score "Cr" (Comprehension Part of the Rate Test)	.408 $\pm$.022
Score "Cs" (Comprehension Test)	.671 $\pm$.014
Score "Ct" (Total comprehension score)	.703 $\pm$.013
Number of Cases	664

parts of the test showed approximately the same reliability as that obtained in earlier comparisons. The coefficients, based on 664 cases, should be thoroughly dependable.

Relative difficulty.- Data from the same distributions also show the extent to which one form of the complete test battery is equivalent to another form. The means and standard deviations for the test as a whole and for each part are given in Table XXIV. The difference between the means for the composite score is only .5 and between the standard deviations .2. In nearly all cases the differences are very small. The greatest difference is in the mean rate scores, a difference of two-tenths of a word per second or about two lines in forty.

TABLE XXIV

RELATIVE DIFFICULTY OF TWO FORMS OF A READING TEST

Score	Number of Cases	Forms 1, B, and 200		Forms 4, A, and 300	
		Mean	σ	Mean	σ
Composite Score	664	117.9	20.5	117.4	20.7
Rate	664	4.03	.94	3.83	.91
Vocabulary	664	30.0	9.1	29.8	10.0
Score "Cr"	664	7.4	1.9	8.2	1.7
Score "Cs"	664	33.0	5.3	33.2	5.4
Score "Ct"	664	47.7	8.0	49.5	7.6

Objectivity of the test.- Each score on this battery of silent reading tests is determined objectively. The scorer does not use his discretion in evaluating any of the student's responses. The rate score is a word count; all other responses are definitely right or wrong.

Standards.- The test is not completely standardized. Only the tentative standards developed during the course of its construction are now available. The means and standard deviations presented in the foregoing paragraph and in earlier sections of this chapter suggest the general level of achievement to be expected. Since many of these averages were based on the scores of the students with whom this study is concerned, they provide approximate standards for these students - that is, for Chicago freshmen. Definite norms based on the scores of an adequate number of unselected students must be established before the test can be equally valuable to other investigators. The meaning of the tentative standards thus far developed will be discussed in the following chapter where, among other topics, the reading habits of the typical Chicago freshman are described.

Practical utility.- Later phases of the investigation demonstrate the value of these reading tests more effectively than any statement which could be given here. Application of the test in the discovery, diagnosis, and correction of reading difficulties is the best measure of its worth.

Summary.- This battery of silent reading tests was designed as a relatively simple and practical measuring instrument which could be used in survey-

ing the general reading ability of a group and in the partial diagnosis of the difficulties of an individual. It purports to be a measure of reading rate, word knowledge, ability to comprehend, and of general reading ability. No claim is made that the test is finally standardized. The standards of achievement thus far established are merely tentative. Furthermore, each part of the test could doubtless be improved to some extent by further refinement, particularly the Test of Accurate Interpretation in Reading. In spite of these obvious limitations the test battery is a useful measuring instrument. Four test forms approximately equal in difficulty have been developed. Each test item has been carefully evaluated. The test battery as a whole shows a reliability of almost .90 and the reliability of each part lies within the range of .70 to .90. The tentative standards now available are helpful in interpreting test scores. And finally, from an administrative point of view, the test is very practical. It is entirely objective, easy to administer, and economical of the time of students and examiners.

CHAPTER III

THE READING ABILITY OF COLLEGE FRESHMEN

The reading tests described in Chapter II were used in carrying out the remedial program which formed the final step in this investigation. Chapter IV reports that project. Before the remedial instruction was undertaken, however, an analysis of the reading achievements and reading habits of Chicago freshmen was begun. Scores on the newly developed reading tests provided certain units of measurement in terms of which reading ability could be analyzed and interpreted. These test scores were supplemented by the eye-movement records of fifty students.

This chapter discusses the reading ability of Chicago freshmen as revealed by the use of the two kinds of measurement mentioned above. It begins with a discussion of the silent reading achievements of college freshmen, expressed in terms of reading test scores. Data are presented relating to (1) the typical achievement of college freshmen, (2) individual differences in reading ability, (3) minimum standards for college students on the tests developed in this study, (4) the inter-relationship of specific abilities, and (5) the extent to which serious reading difficulties are found among entering freshmen. Later sections of the chapter deal with the laboratory study, discussing the reading habits and achievements of fifty students as indicated by eye-movement records.

Typical Achievement as Revealed by Test Results

The normal reading ability of Chicago freshmen is accurately revealed by the scores of the 664 freshmen in the 1931-32 freshman class. Therefore, their scores on Form 1 of the rate test, Form B of the vocabulary test, and Form 200 of the comprehension test will be discussed at some length. These distributions are as representative and as comprehensive as any which were obtained in this investigation.

Normal rate of reading.- The distribution of the rate scores of 664 college freshmen appears in Diagram 2. The histogram is fairly symmetrical, but a slight positive skewness may be observed. That is, there was less variability in rate among

readers below the average than among those above the average. Approximately two-thirds of the students' scores (those lying less than one sigma above and below the mean) were within the range, 3.00 to 5.00. Therefore, in so far as these data are dependable, rate scores anywhere between 3.00 and 5.00 represent normal performance. To be sure, there is a vast difference between a reading rate of three words per second and one of five words per second - a difference of two pages out

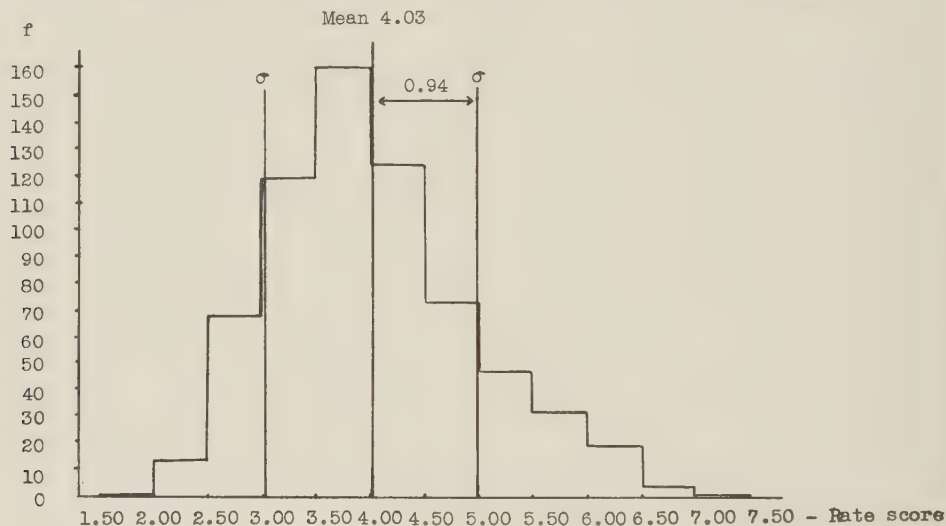


Diagram 2.- Distribution of the rate scores of 664 Chicago freshmen on Form 1 of the Test of Achievement in Silent Reading.

of every five. Nevertheless, that range in reading rate was found among students in the middle and most homogeneous section of this class.

The mean for the group (4.03) indicates that the average Chicago freshman reads approximately four words per second. In other words, he devotes three minutes to the reading of two ordinary pages in a college textbook.¹ At that rate he can read approximately 40 pages in one hour. This standard is not a rigid one, for many factors operate to make reading rates vary. A number of such factors

¹A rate of 4.00 means 240 words per minute or 720 words in three minutes. Ordinary textbook pages with four-inch lines carry at least 350 words. Therefore, three minutes to two pages seems a fair estimate.

have been pointed out in other parts of this report. In addition, the mechanical features of books must be taken into account. Reading rate is influenced by factors such as: size of type, the number of variations in size and style of type, use of tables, charts, and diagrams, the headings, indentations, spacing, length of line, and number of lines on each page. In spite of the differences in rate which must be expected, the average rate as stated above should prove suggestive to college teachers.

One precaution seems advisable; namely, that average rate among college students today is not a safe standard for the future. The question is frequently asked, How fast should college students read? To that question, the average reading rate of college students today provides no final answer, even though it is the best single standard now available. The average reader at the present time may be quite inefficient; perhaps he could double his rate of reading. The fact that a few readers are able to read understandingly at the rate of 7.00 or more words per second suggests that possibility. Obviously, optimum rate is maximum rate with maximum comprehension. Therefore, instead of setting up average rate as a standard, the teacher should encourage each student to build up his reading rate to the point where his own optimum efficiency is realized. In some cases, the present average rate may be as high a standard as the student should attempt; in others, it may be much too low.

Vocabulary.- The vocabulary test scores of 664 college freshmen are summarized in Table XXV. The average score on the test was 30. In other words, the average Chicago freshman can identify the synonym for about three-fifths of the words in the Test of Reading Vocabulary. Only about one student in seven is familiar with fewer than 20 of the 50 test words. Note that 451 of the 664 scores were between 20 and 40.

Comprehension.- Table XXVI presents a frequency distribution of the total comprehension scores ("Ct") of the class. The possible range is 0 to 70; the actual range in this instance was 20 to 66. The mean for the group was approxi-

TABLE XXV

DISTRIBUTION OF THE SCORES OF 664 CHICAGO FRESHMEN ON
FORM B OF THE VOCABULARY TEST

Interval	f	
45 - 50	24	
40 - 45	93	
35 - 40	116	
30 - 35	114	Mean = 30.0
25 - 30	123	
20 - 25	98	
15 - 20	69	
10 - 15	24	$\sigma = 9.1$
5 - 10	3	

mately 48, with the middle two-thirds of the scores falling between 40 and 56. Scores approximating 50 are therefore typical for Chicago freshmen.

TABLE XXVI

DISTRIBUTION OF THE COMPREHENSION SCORES OF 664
CHICAGO FRESHMEN

Interval	f	Interval	f	Interval	f
64 - 66	3	48 - 50	70	32 - 34	18
62 - 64	10	46 - 48	51	30 - 32	7
60 - 62	19	44 - 46	50	28 - 30	8
58 - 60	20	42 - 44	40	26 - 28	1
56 - 58	51	40 - 42	37	24 - 26	3
54 - 56	60	38 - 40	31	22 - 24	1
52 - 54	77	36 - 38	21	20 - 22	1
50 - 52	68	34 - 36	17		
Mean					47.7
Standard Deviation					8.0

Composite score.- The composite scores on the test battery as a whole are represented by Diagram 3. Theoretically, a maximum score of 200 would be possible. In this instance no score was higher than 170. The average, 117.9, is almost the exact sum of the three separate averages already given - 40 (ten times the average rate score, 4.00) plus 30 (the average vocabulary score) plus 48 (the average

comprehension score). The distribution of the composite scores is quite symmetrical. Applying the formula¹ $S_k = \frac{3 (M - Md)}{\sigma}$ as a measure of the skewness of the distribution, S_k has a value of $-.02$, indicating a very slight negative skewness. Two-thirds of the scores fall within the general limits, 100 to 140.

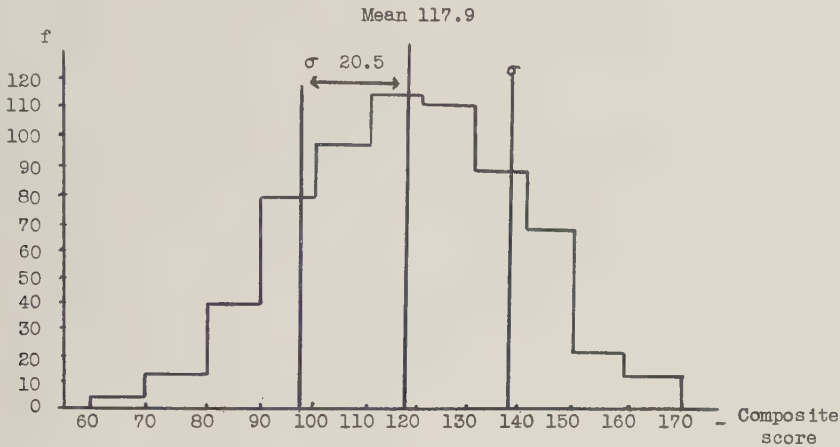


Diagram 3.- Distribution of the composite scores of 664 Chicago freshmen.

Interpretation of these standards.- Typical achievement as described in the foregoing paragraphs is the typical achievement of freshmen in a college where entrance standards are relatively high. A Chicago freshman class is by no means an unselected group. Most of the students admitted rank high in their respective high-school classes and measure up well on all entrance tests. The standards given above satisfy the purposes of the present investigation - for Chicago freshmen must meet Chicago standards. They are also applicable in any college where similar methods of selection are employed. When the tests are used with unselected groups, somewhat lower averages and perhaps greater dispersion should be expected.

¹See Karl J. Holzinger, Statistical Methods for Students in Education, p. 123. Boston: Ginn and Co., 1928. The median which must be substituted in the example above is 118.3; the mean and standard deviation are given in Diagram 3.

Individual Differences as Shown by Test Scores

As early as 1897, Quantz¹ reported that there were important differences in the reading ability of college students. Later investigators verified and extended his observations, especially Anderson, whose study was focused on this question.² The test scores obtained in the present study substantiate the findings of previous investigators, but since they constitute new units of measurement and portray the existing differences in clearly defined terms, the following data are presented.

Individual differences in general achievement.- The composite scores of two of the 664 Chicago freshmen were lower than 65. At the other extreme was one score of 170. The former was about two and one-half standard deviations below the mean; the latter, two and one-half standard deviations above it. These data indicate extreme differences in silent reading ability among mature readers. If differences are so pronounced in this group of students, carefully selected on the basis of academic attainments, they are undoubtedly greater among mature readers generally. The fact that reading is a tool subject and that reading habits are formed in the elementary school should not lead one to assume that mature readers are equally competent, or even approximately so.

Individual differences in reading rate.- One student made a rate score lower than two words per second (2.00). Two students made rate scores higher than seven words per second (7.00). This difference of five words per second means that the rapid reader can read 18,000 more words in one hour than can the slow reader. Allowing 400 words to the page, the rapid reader can read 45 pages per hour more than the slow reader; that is, the latter can read only 18 pages while the former is reading 63. Stated in another way, it would take the slow reader ten and one-half hours to complete the same reading assignment that the rapid reader could complete in three hours. If the comprehension of the rapid reader is equally good, his advantage is tremendous.

¹J. O. Quantz, "Psychology of Reading," Psychological Review, Series of Monograph Supplements, Volume II, Number 1 (Whole Number 5), December, 1897.

²Ernest M. Anderson, Individual Differences in the Reading Ability of College Students. University of Missouri Bulletin, XXIX (1928), Number 39.

Differences in the ability to recognize word meanings.- Three of the 664 students made vocabulary scores lower than 10. Twenty-four students made scores higher than 45. The former students attach no meaning, or an incorrect meaning, to about four-fifths of the words in the test. They know relatively few of the words in the "intelligent readers' vocabulary." They can read only simple material with adequate comprehension. If this test is representative, they need to consult the dictionary about eight times as frequently as their better informed classmates in order to learn the meaning of rather common, non-technical words.

Differences in comprehension.- Similar differences in comprehension exist. The lowest of the 664 comprehension scores ("Ct") was 20; the highest was 66. That is, the best reader with respect to comprehension could respond correctly to more than three times as many of the test items as the most superficial reader in the class. In terms of standard deviations, the lowest score was about three sigmas below the mean; the highest, more than two standard deviations above the mean. This difference is tremendously important. It deals with the results achieved in reading. Assuming, in any given case, that failure to comprehend is due to the manner in which the reading is done, the poor reader must re-read a passage in order to understand it as well as his more fortunate classmates. Inadequate or inaccurate interpretation corrected by re-reading may result in as much or more loss of time in study than an abnormally slow rate of reading. Inadequate comprehension uncorrected defeats the reader's purpose.

Minimum Standards on the Reading Tests

The test scores presented in the preceding paragraphs provide minimum standards of performance for Chicago freshmen on the reading tests developed in this investigation, that is, certain points in the several distributions which represent danger points. These cannot be rigidly applied. One student will compensate by special effort for poor reading habits which cause other students to fail in college. Nevertheless, they are significant, not only to those who direct the reading activities of Chicago freshmen but also, to some extent, to college

teachers generally.

Minimum rate.- A rate score lower than 3.00 is a definite handicap to the college student. Exceptionally good comprehension combined with an unusual amount of perseverance may bring the student success, but in many cases, students who read more slowly than three words per second become hopelessly discouraged and withdraw from college.¹ The demands of the college weigh heavily when day after day a student who reads at the rate of 2.00 or 2.50 words per second is confronted with long reading assignments. A point score of 3.00 is approximately one standard deviation below the mean for Chicago freshmen. In other words, only about 17 per cent of the students in the 1931-32 class read more slowly than three words per second.²

Other minimum standards.- Points located approximately one standard deviation below the mean for Chicago freshmen provide the best minimum standards now available for the vocabulary score, the comprehension score, and the composite score. In colleges where entrance standards are lower, somewhat lower minimum standards might be acceptable. For Chicago freshmen, scores lower than one standard deviation below the mean indicate that the student will work under a distinct handicap in that institution. The following minimum standards (derived as implied above) seem approximately correct: vocabulary score, 20; comprehension score "Ct", 40; composite score, 90.

Relation between Rate and Comprehension

From time to time studies have appeared purporting to show the relation between rate and comprehension in silent reading. The evidence is contradictory.³ In 1916 King reported that he found negative correlations between rate and compre-

¹See page 77.

²Apparently this standard is conservative. Mrs. L. C. Pressey, on the basis of her extensive investigations, advocates a minimum standard of 200 words per minute or 3.33 words per second. See Mrs. L. C. Pressey, Manual of Reading Exercises for Freshmen, p. 3. Columbus, Ohio: Ohio State University Press, 1928.

³Irving King, "A Comparison of Slow and Rapid Reading," and "A Comparison of the Efficiency of Slow and Rapid Readers," School and Society, IV (1916), 830-34, and VI (1917), 203-04.

hension test scores. The coefficients varied from $-.07$ to $-.47$. Later studies by ¹Eurich, ²Gray, ³Judd, ⁴Sister Kathleen, and ⁵Witty and Lehman report a positive relationship. Witty and Lehman obtained coefficients of $.78$, $.81$, and $.86$ between comprehension as measured by the Thorndike-McCall test and rate as measured by their own rate test. Although all the later studies report positive correlation between rate and comprehension, the coefficients vary greatly. Eurich reports 26 coefficients derived from the scores of various rate and comprehension tests ranging from $.04$ to $.67$. The average of the 26 coefficients is $.31$. He concludes that the relationship between rate and comprehension is "positive but not close" and that it is "dependent upon the manner in which each is measured."

The data in Table XXVII show the extent to which rate and comprehension are related when measured by the tests described in Chapter II. The first coefficient was obtained by correlating the rate score of 172 college freshmen on Form 1 of the rate test with the sum of their "Cr" scores on Forms 1 and 2 of the test. The second and third were based on test scores from a single form of the rate test. The last was obtained by correlating rate with the score "Ct" which is derived in part from the Test of Achievement in Silent Reading and in part from the Test of Accurate Interpretation in Reading. The four coefficients are positive but low.

¹Alvin C. Eurich, "The Relation of Speed of Reading to Comprehension," School and Society, XXXII (September, 1930), 404-06.

²William S. Gray, Studies of Elementary School Reading through Standardized Tests, p. 134. Supplementary Educational Monographs, Vol. I, No. 1. Chicago: the University of Chicago, 1917.

³Charles H. Judd, Measuring the Work of the Public Schools, p. 155. Survey Committee of the Cleveland Foundation, 1916.

⁴Sister M. Kathleen, "A Comparative Study of Certain Silent Reading Tests," Catholic Educational Review, XXII (1924), 589-95; XXIII (1925), 33-42.

⁵P. A. Witty, and H. C. Lehman, "Teaching the College Student How to Study," Education, XLVIII (September, 1927), 47-56.

They tend to support Burich's conclusion that the relation between rate and comprehension is positive but not close.¹

TABLE XXVII
CORRELATION BETWEEN RATE AND COMPREHENSION

Test Scores Correlated	Number of Cases	r ± P. E.
Rate Form 1 with Score "Cr" Form 1 plus Score "Cr" Form 2	172	.127 ± .051
Rate Form 1 with Score "Cr" Form 1	291	.104 ± .039
Rate Form 2 with Score "Cr" Form 2	300	.126 ± .038
Rate Form 1 with Score "Ct", Forms 1 and 200	664	.270 ± .024

Other Interrelationships

Intelligence and reading ability.- In his study of "Probable Types of Difficulty Affecting Comprehension in Reading," Hilliard mentions intelligence as the most important single factor.² Other studies - especially at the elementary-school level - have dealt with the relation between intelligence and reading ability.

¹Many of the coefficients which other investigators have reported dealing with the relation between rate and comprehension in reading are spuriously high because both scores used in the computations measure the same thing. The Monroe Silent Reading Test may be cited as an illustration. The comprehension score depends largely on the rapidity of the reader. A slow reader cannot possibly make a high comprehension score on the Monroe test. Consequently, high correlations result, not from the relation between speed and comprehension, but because both scores on the test measure the same ability. Test scores from any timed test of comprehension such as the Burgess, the Chapman-Cook, the Monroe, or the Whipple High School and College Reading Test should correlate highly with rate scores. Such coefficients are apt to be decidedly misleading. For example, although Miles A. Tinker "The Relation of Speed to Comprehension in Reading," School and Society XXXVI (July 30, 1932), 158-60 rightly points out the need for measuring rate and comprehension in the same or strictly comparable situations, he entirely overlooks the fallacy mentioned above and concludes that speed of reading consistently shows a high correlation with comprehension. For a discussion of the question from various angles see, William S. Gray, Summary of Investigations Relating to Reading, pp. 124-31. Supplementary Educational Monographs, Number 28. Chicago: Department of Education, University of Chicago, 1925.

²G. H. Hilliard, Probable Types of Difficulty Affecting Comprehension in Reading, University of Iowa Studies in Education, Vol. II, Number 6. Iowa City: the University of Iowa, 1918.

Positive correlations are reported. Anderson¹ found a close correlation between intelligence and reading ability among college students, especially between intelligence and word knowledge. He pointed out, however, that since the intelligence test score was doubtless influenced by reading ability, the observed relationship was probably closer than the true relationship. The data which follow agree essentially with the findings of previous investigators.

The correlation between intelligence and reading ability is suggested by the coefficients in Table XXVIII. The percentile rank of the student on the American Council Psychological Examination was used as a measure of intelligence. The reading scores were derived from the tests developed in this study. The scores of the first 200 students on the alphabetical list of 664 University of Chicago freshmen were used in the calculations. The first coefficient indicates that general reading ability and intelligence are closely related. A similar relationship has been reported when other reading tests and other measures of intelligence have been employed. For example, the authors report a correlation of .827 between the Iowa Silent Reading Test and the Terman Intelligence Test.² If intelligence is an important factor in developing good reading habits, the correlation should be positive and relatively high, but unless intelligence is the one dominant factor, the correlation should not be extremely high. If the relationship between two tests is as high as the reliability of the tests themselves, it would seem that both are measuring the same thing. The first coefficient in Table XXVIII is considerably lower than the reliability of either test yet high enough to indicate that reading and intelligence are closely related.

A low but positive correlation was found between rate of reading and intelligence. The correlation between intelligence and the reading vocabulary was rather high, a conclusion reported frequently by other investigators. The

¹Op. cit.

²H. A. Greene, and A. N. Jorgensen, Examiner's Manual for Iowa Silent Reading Tests, p. 4. Iowa City: Bureau of Educational Research and Service, University of Iowa, 1927.

correlation between intelligence and comprehension was moderately high. To the degree that the comprehension part of this reading test is a true measure of that

TABLE XXVIII
CORRELATION BETWEEN INTELLIGENCE AND READING ABILITY

Correlation between Intelligence and:	Number of Cases	r $\pm$ P. E.
Composite Score	200	.702 $\pm$.024
Reading Rate	200	.312 $\pm$.043
Reading Vocabulary	200	.724 $\pm$.023
Comprehension (Score "Ct")	200	.569 $\pm$.033

ability, it seems evident that certain elements other than intelligence are important factors in producing adequate comprehension.

Word knowledge as a factor in reading.- Although both the rate and comprehension parts of the reading test are supposedly free from technical or especially difficult words, there appears to be a relationship between the reading vocabulary and reading rate and between vocabulary and comprehension. The vocabulary scores of the 664 Chicago freshmen showed a correlation of .380 $\pm$.023 with reading rate. Since the vocabulary test is practically untimed, it seems unlikely that the observed relationship was due to speed of reading on the vocabulary test to any considerable degree, but rather to a reading rate developed in part on the basis of word knowledge.¹

The correlation between the same set of vocabulary scores and the students' comprehension scores (the score "Ct") was .607 $\pm$.017. This coefficient may be somewhat too high because both the vocabulary test and the comprehension test correlate

¹Traxler found a somewhat higher correlation between reading rate and word knowledge at the junior-high-school level. Arthur E. Traxler, The Measurement and Improvement of Silent Reading at the Junior-High-School Level, p. 94, Doctor's Dissertation, Department of Education, the University of Chicago, 1932 (Private edition).

well with intelligence. However, since the path of influence leads in both directions, the elimination of intelligence by the method of partial correlations gives a coefficient that is too low. When that method is applied, the correlation between vocabulary and comprehension is $.343 \pm .024$.¹ The true value lies between the zero- and first-order coefficients, .607 and .343. An exact coefficient is of less importance than the fact of positive relationship which these data imply.

Reading ability and academic achievement.- At the end of the winter quarter 1931, a study was made of the relation between reading ability and academic achievement among freshmen at the University of Chicago. The former was based on reading test scores; the latter on grade-point records for two quarters. The Test of Achievement in Silent Reading had been given as one of the entrance tests in October, 1930. In January, 1931, the Minnesota Reading Examination had been administered to 255 freshmen, a group selected on the basis of test scores on the former test containing 86 superior readers, 61 poor readers, and 108 doubtful cases, that is, students who seemed likely to be in need of remedial work in reading. This sample represented both extremes in reading ability. It was probably weighted with students whose reading ability was below the average for the class as a whole. Nevertheless, it contained a number of students of average ability. Correlations based on these cases should not be spuriously high because of the influence of selection. At the end of the winter quarter there were 168 students who had comparable records on the two reading tests mentioned above and also grade-point records based on a uniform number of majors completed during two quarters of residence work.

The coefficients presented in Table XXIX indicate the relation between

¹The formula used in this study in computing partial correlations with three variables (in this problem vocabulary, comprehension, and intelligence) is given by Holzinger, op. cit., p. 287:

$$r_{12.3} = \frac{r_{12} - r_{13}r_{23}}{(1 - r_{13}^2)(1 - r_{23}^2)}$$

The zero-order coefficients substituted in this instance were: $r_{12} = .607$, $r_{13} = .724$ (see Table XXVIII), and $r_{23} = .569$.

reading ability and academic achievement. A correlation coefficient between intelligence and achievement is also included in the table. Each coefficient obtained was positive and significant. The lowest one was five times its probable error. Intelligence has often been emphasized as an essential element in predicting academic success, yet among these coefficients, the one based on intelligence and achievement is little higher than those based on reading ability and achievement. The argument has been advanced that the apparent relationship be-

TABLE XXIX
CORRELATION BETWEEN READING TEST SCORES AND NUMBER
OF GRADE POINTS OF 168 UNIVERSITY FRESHMEN

Correlation between Grade Points and:	$r \pm P. E.$
Reading Rate	.240 $\pm$.048
Comprehension (Score "Cr")	.323 $\pm$.046
Minnesota Total Score	.375 $\pm$.044
Minnesota Comprehension	.351 $\pm$.045
Minnesota Vocabulary	.313 $\pm$.047
Intelligence (Percentile)	.443 $\pm$.042

tween reading and achievement is due to their common relationship to intelligence. When the method of partial correlations is employed, holding intelligence constant, the two lowest coefficients are still positive and about three times their probable errors. The correlation between rate and grade points with intelligence constant, is .163 $\pm$.05 and that between the comprehension score "Cr" and grade points, .140 $\pm$.051.¹ Again the true value probably lies somewhere between the zero-order and first-order coefficients since the path of influence

¹In these computations the zero-order coefficients were: (1) $r_{12} = .240$, $r_{13} = .220$, and $r_{23} = .443$; (2) $r_{12} = .323$, $r_{13} = .480$, and $r_{23} = .443$. The first and third coefficient for each problem is given in Table XXIX. The second coefficient in each series was computed for the scores of these students by Karl J. Holzinger for a report to the Dean of the College, University of Chicago, entitled "Preliminary Report on the Application of the Sones-Harry and Other Tests to University Freshmen." Unpublished report, November 1930.

between intelligence and reading ability leads in both directions. Although the coefficients are not high, they indicate a relationship that is probably as close as that between achievement and most other single factors.

As a further means of determining the relationship between reading ability and success in college, an analysis was made of the records of the 47 freshmen who had withdrawn from the University prior to the end of the winter quarter 1931. In Table XXX the students are classified into groups with respect to reading ability and intelligence. In this classification those whose reading scores were above average on all the reading tests constitute the high group, those whose scores were below the lower quartile on all reading tests the low group, and all other students the average group. Many of the latter are below average in reading ability as the term "average" is ordinarily used. With respect to intelligence, the low group is made up of students whose percentile ranks fell between 1 and 24; the average group, of students ranking from 25 to 74; the high group, of students with a percentile rank higher than 75. Sixteen of the 47 students had been dismissed outright because of poor work and in 10 cases poor work was so clearly in evidence that it was, without doubt, the major cause of withdrawal. This fact should be noted in connection with Table XXX. In other words, reading handicaps and low intelligence leading

TABLE XXX

CLASSIFICATION OF 47 STUDENTS WHO WITHDREW FROM THE
UNIVERSITY BEFORE THE END OF TWO QUARTERS

Intelligence Reading	No Data	Low	Average	High	Total
High	---	---	2	3	5
Average	2	11	7	2	22
Low	1	16	3	---	20
Total	3	27	12	5	47

to discouragement and failure in college work accounted for a considerable proportion of the withdrawals.

Complete data concerning the achievements of the 1931 class are not available. However, the following facts are worth consideration. On the basis of the entrance test scores, 95 of the 664 students were selected as those most likely to be handicapped in reading. The following data from the Recorder's office show the quality of work which they did during the autumn quarter:¹

- 1 student had withdrawn
- 18 students received "s" in not a single course
- 25 students received "s" in only one course
- 22 students received "s" in two courses
- 23 students received "s" in three courses
- 6 students received "s" in four courses

Conversely:

- Only 19 students received no "u" or "r" report
- 25 students received one "u" or "r" report
- 26 students received two "u" or "r" reports
- 19 students received three "u" or "r" reports
- 5 students received four "u" or "r" reports

The total number of "s" reports was 162. The total number of "u" and "r" reports was 154. In other words, during the autumn quarter the 95 poorer readers in the 1931 class did satisfactory work in scarcely more than half their classes.

Extent of Reading Difficulties

Since a freshman class at the University of Chicago is a selected group, reading disabilities are doubtless present in fewer cases than in an equal number of unselected college freshmen. Therefore, whatever evidence there is that reading disabilities are present among Chicago freshmen is evidence that similar disabilities exist in even greater proportions in many other colleges.

Among the 643 freshmen in the entering class in October, 1930, there were 65 students who, according to test scores, were distinctly handicapped with respect to reading ability. Their rate of reading was below an acceptable standard and their vocabulary and comprehension scores were also low. In addition, there were many students who were low in either rate, vocabulary, or comprehension. In other words, about 10 per cent of the students were inefficient readers by every criterion

¹The grading system was as follows:

"s" = satisfactory;

"u" = unsatisfactory;

"r" = registered, - insufficient evidence for "s" or "u."

available and a number of others were handicapped by specific disabilities.

A somewhat more accurate description is afforded by the test scores of the 664 freshmen in the 1931 class. Eighty-two students had rate scores lower than 3.00 on Form 1 of the rate test. Ninety-six had rate scores lower than 3.00 on Form 2. Fifty had rate scores lower than 3.00 on both Forms 1 and 2. Eighty-one students had scores lower than 20 on both Forms A and B of the vocabulary test. Fifty-one students made low comprehension scores on both forms of the comprehension test - more than one sigma below the mean. In fifteen cases, students' scores were lower than the foregoing standards on rate, comprehension, and vocabulary on both forms of the test. On the basis of composite scores, eighty-one students would be classified as poor readers.¹

In the class of 1930 and in that for 1931, serious reading disabilities were discovered in 10 to 12 per cent of the cases. In many other instances a particular phase of reading ability was underdeveloped. Approximately 20 per cent of the students made unsatisfactory scores on one or more parts of the test battery. Since reading disabilities seriously handicap so many students, the need for a satisfactory method of group instruction in reading becomes apparent.

Before describing the remedial program developed in this study for use with students in the 1930-31 class, however, an analysis of reading ability based on eye-movement records will be presented.

The Significance of the Laboratory Study

Eye-movements have long been recognized as important symptoms of reading ability, and numerous studies have been made of eye-movement records. Studies by Dearborn, C. T. Gray, Judd, Buswell, and others have done much to develop the principles upon which instruction in reading should be based. Other investigators have studied special problems such as the reading of numerals, formulae, and foreign languages, the effect of peripheral vision in reading, eye-movements in learning

¹Students whose composite scores on both forms of the reading test were more than one sigma below the mean.

to spell a word, and eye-movements in reading music. In none of the previous studies, however, has special reference been made to the ordinary reading habits of mature readers. College students have served as subjects in certain instances, but in previous investigations emphasis has not been placed on the reading performance of college students when confronted with materials drawn from their own textbooks. Previous studies have shown that, in general, the mature reader makes comparatively few fixations per line, that each fixation is brief, and that progress along the line is regular and rhythmic.¹ They have also demonstrated that the quality of the material and the purpose of the reader affect his manner of reading.² However, since definite facts about the eye-movements of college students when reading paragraphs similar to those they encounter in their daily work have not been reported, the laboratory work described in the following pages constitutes a new approach in analyzing their reading habits.

The specific purposes of the laboratory study as presented in Chapter I need not be repeated. The general procedure has also been outlined, but at this point a fuller description of the technique seems in place.

Description of Laboratory Procedure

Selection of cases.- Two groups of students were selected for the experiment. One group was composed of 25 of the best readers who could be discovered in the 1930 freshman class. The other was made up of 25 of the most inefficient readers in the class. Test data were used to identify the good and poor readers. The students' records for the first quarter at the University of Chicago were also considered. The final selection of cases was made through personal interview with the students. All students who participated in the experiment did so voluntarily.

¹G. T. Buswell, Fundamental Reading Habits. Supplementary Educational Monographs, Number 21. Chicago: Department of Education, University of Chicago, 1917.

²Charles H. Judd and Guy T. Buswell, Silent Reading: A Study of the Various Types, Supplementary Educational Monographs, No. 23. Chicago: University of Chicago Press, 1922.

Consequently, the groups include the best and poorest readers who were willing to participate, not necessarily the best and poorest in the class. The groups doubtless contain a few of the best and poorest respectively and the evident contrast in ability will be pointed out presently.

In interpreting the data, it should be remembered that the students represent the two extreme levels of ability in the freshman class of a university where the entrance standards are relatively high. Consequently, although the group of superior readers selected from this class may not be the most efficient readers which could be found in an entering class, in all probability they are, as a group, much better readers than the best in many unselected freshman classes. In the same way, the poor readers whose eye-movements were photographed in this investigation may, or may not, be typical of the poorest readers in other entrance classes.

Reading materials selected.- Seven paragraphs were used in the experiment. Three were adapted from the Test of Achievement in Silent Reading. The last four in the series were adapted from a college textbook.¹

The first, second, third, and fourth paragraphs were intended to be approximately equal in difficulty. The fifth was designed to vary with respect to the difficulty of the vocabulary. A number of words in the original paragraph were replaced by harder synonyms. The sixth paragraph consisted of a single long sentence. It was designed to discover the influence of sentence structure on reading. The final paragraph had a simple vocabulary and was written in simple, straightforward sentences but was selected because the ideas expressed were more abstract. It dealt with "the self" and its meaning seemed to be less tangible than that of the other paragraphs.

Each paragraph was mounted under thin glass, making a slide approximately three by five inches. The paragraphs were printed in 11-point type with lines four inches in length. Paragraph VI had 15 lines; each of the others had 14. A small dot of ink was placed in the upper left-hand corner and in the upper right-hand

¹R. F. Piper, and P. W. Ward, The Fields and Methods of Knowledge. New York: Alfred A. Knopf, 1929.

corner of each slide directly above the first and last letters in full-length lines of print. Slides of similar size and appearance were prepared, bearing comprehension exercises based on the paragraphs.

Procedure in photographing.- Reference was made in Chapter I to printed descriptions of the apparatus and its use. The general procedure in this experiment was the same as that in the reports cited. The following details should be noted.

As soon as the subject was properly seated before the camera a slide was placed on the reading plate bearing the following information:

These are your directions for reading. You will be told to "Look at the left dot," then "Look at the right dot," then "Read." When you finish reading a paragraph close your eyes. Try to read just the way you usually do when you read a book or newspaper. Forget about the camera and think about what you are reading. Read as rapidly as you can understand the story.

The subject was told to read this slide at pleasure while the camera was being focused. As soon as the camera was in focus, another slide bearing the following specific directions was placed on the reading plate. The experimenter read the directions orally as the subject read them silently:

Read this paragraph silently just as you normally read when you wish merely to understand what a paragraph is about. Read it as rapidly as you are able to grasp the essential ideas. Do not try to remember details, but just read to understand what the paragraph is about.

Paragraph I was then presented as though a photograph were being made. As a matter of fact, the light was focused just off the film. The motor was started and allowed to run a few seconds in order to dispose of the piece of film which is always exposed in loading or focusing the camera. When the subject had finished reading the paragraph, the experimenter removed it and slipped the first comprehension slide into place on the reading plate, saying as he did so, "Which of the following statements expresses the central thought of the paragraph?" While the subject was making his decision, the focus of the camera was tested and the camera was shifted so that the rays of light which produce the record would be recorded on the film. As soon as the subject gave his response, it was recorded on a blank prepared for that purpose and the experimenter said, "Are you perfectly satisfied

with your choice? If you have any doubt, you may express a second choice."

The procedure for Paragraph II was identical, except that in this instance a photograph was made.

Paragraph III was preceded by the following directions:

Read this paragraph thoughtfully and critically. Make a vigorous effort not only to find out what it is about but also to grasp all the important ideas. Do not pause nor reread unless it is absolutely necessary. Proceed as you usually do in careful, thoughtful reading.

The comprehension slide carried the usual list of statements, but this time instead of selecting the central thought of the paragraph the subject was asked to state, "Which of the following statements are true? Just say, 'The first is true (or false); the second is true (or false); etc.' Skip around if you like." In addition the student was asked to give in his own words three definitions of success suggested in the paragraph and to state which of the three the author proposed as the best one.

The procedure for Paragraphs IV, V, and VI was identical with that for Paragraphs I and II. The same directions were also used to introduce Paragraph VII but true-false comprehension exercises were used with the final paragraph.

The only unique feature in the foregoing procedure is the attempt to standardize the reader's response.

Evaluation of the technique.- Using the first paragraph as a trial selection proved to be a very helpful device. Not only did the reader become accustomed to the laboratory situation, but he demonstrated objectively whether or not he understood the directions. Several students attempted to read too rapidly when confronted with Paragraph I. When the simple comprehension exercise was submitted, some remarked that they had really failed to read. From that time on they knew what to expect.

The comprehension technique seemed to serve its intended purpose. It was intended merely to force the subject to read intensively enough to understand the thought of the paragraph. Speed of reading was emphasized in the directions to such a degree that a check on comprehension was indispensable; yet the subject was not

expected to remember details.

Presenting the directions in printed form helped to standardize the procedure and doubtless added to the validity of the results obtained.

Analysis of data.- In analyzing the films, all computations were based on ten lines, disregarding the first three and last line (or lines) in each paragraph. These lines were omitted in order that the student might settle into a definite style of reading on any given paragraph and in order to eliminate differences due to indentation or to incomplete lines. Computations were also greatly simplified, since each total for ten lines with one place pointed off became the average for one line.

The films were projected and the following data recorded: (1) number of fixations per line, (2) number of regressive movements per line, and (3) the time required to read the ten lines expressed in twenty-fifths of a second.¹ The average duration of fixations was then determined by dividing the total reading time² for the ten lines by the number of fixations made in reading them.

A comprehension score was assigned to each student based on his performance when reading before the eye-movement camera. It was computed as follows: (1) a credit of five points was allowed, in the case of Paragraphs I, II, IV, V, and VI, if the student gave the correct answer as his first choice, (2) a credit of two points was allowed with these paragraphs if he gave the correct answer as his second choice, and (3) ten points were allowed in the case of both Paragraph III and Paragraph VII, one point for each item in the exercise. If all exercises were done perfectly, the score was 45.

¹Tables XL and XLI, p. 96, show how the data were recorded.

²This method involves a slight constant error, since it disregards the time required to shift the eye from one fixation point to the next. Each average is a trifle larger than it would be if it were based solely on the cumulation of the several fixations. However, since the movement of the eye is extremely rapid - too rapid for vision during the transition - this process occupies an insignificant part of the total reading time and averages computed on the basis of total reading time are essentially correct.

Presentation and Interpretation of the Data

A comparison of the two groups of students.- In the discussion which follows, the 25 superior readers will be referred to as Group I; the 25 poor readers, as Group II. Individuals will be identified by assigned numbers.

Table XXXI is presented to indicate the difference in ability between Group I and Group II as revealed by certain test scores. Only three students in Group

TABLE XXXI
TEST SCORES OF GROUP I AND GROUP II

Group I					Group II				
Student Number	Intelligence Percentile	Rate Score	Minn.* Compr.	Minn.* Vocab.	Student Number	Intelligence Percentile	Rate Score	Minn.* Compr.	Minn.* Vocab.
27	97	4.83	72	82	56	10	1.66	42	22
7	86	5.08	72	46	51	35	2.08	63	20
14	97	4.56	84	66	53	02	2.48	39	18
26	..	7.00	..	..	60	28	2.23	60	24
1	..	7.00	69	81	12	06	2.14	45	16
19	57	4.41	66	54	63	28	2.58	42	23
22	90	6.67	63	54	32	57	2.69	45	31
25	76	7.00	63	58	54	57	2.69	39	21
29	28	4.83	69	45	31	..	3.46	27	17
34	90	4.16	78	47	58	..		48	36
10	70	4.90	75	55	38	35	2.48	51	24
28	22	3.47	48	44	47	17	3.32	45	26
49	99	5.08	51	60	50	01	2.57	33	15
62	28	5.25	48	28	3	86	3.95	72	87
16	94	6.48	63	57	55	35	2.57	51	21
15	..		69	42	52	35	2.73	51	28
8	96	5.90	78	60	57	..		..	..
17	94	5.53	66	45	48	06	3.73	45	24
33	86	5.90	72	52	59	04		30	10
5	70	6.30	72	50	61	17	4.08	57	30
4	99	4.99	75	56	39	22	2.88	51	26
41	76	5.53	66	42	37	10	3.82	45	18
6	86	4.00	69	48	30	43	2.48	48	33
46	..	4.42	42	60	40	06	2.89	54	19
42	63	4.83	57	66	36	17	2.41	69	24
Mean	76.4	5.34	66.1	54.1	Mean	25.3	2.81	48	25.5
Average for the Entire Freshman Class						50	3.68	54.4	35.2

* The scores entered in this column are raw scores - not percentile ranks.

I had percentile ranks lower than 50 on the intelligence test given at the time of entrance. Only three students in Group II ranked higher than 50. Note the contrast expressed by the means: 76.4 in the case of Group I and 25.3 for Group II. The relation between reading ability and intelligence mentioned earlier in this chapter is reaffirmed by these data. The difference between Groups I and II with respect to reading ability is equally pronounced. Only one student in Group I read more slowly than four words per second. In Group II only one student read as rapidly as four words per second. The average rate for Group I was 90 per cent faster than that for Group II. Only five students in Group I made comprehension scores lower than 60, while in Group II only four students made scores of 60 or better.¹ The average comprehension score for Group I was 11.7 points above the average for the entire freshman class. The average for Group II was 6.4 points below the class average. In vocabulary, the average for Group I was 54.1; for Group II, 25.5. From these data it is evident that the two groups of students represent the extremes in reading ability. There is even little overlapping in the test scores; one group begins approximately where the other leaves off.

Differences in the mechanical reading habits of the two groups.- From the data in Tables XXXII and XXXIII, the groups may be contrasted on the basis of their performance in the laboratory. The totals from the eye-movement record sheets are brought together in these tables. The second column in each table shows the total number of fixations made by each student in reading 60 lines, — ten lines in each of six paragraphs. The students are so listed that the total number of fixations appear in ascending order in Table XXXII; in descending order in Table XXXIII.² Other columns in the tables summarize the data with respect to regressive movements, total reading time, and adequacy of comprehension.

¹Note that these are based on the Minnesota test - not on the Test of Accurate Interpretation in Reading.

²The students were listed in the same order in Table XXXI, and will appear in that order in all the tables which follow.

The tables are to be interpreted as follows: Whereas the student, Number 27, read 60 lines with 266 fixations and only 32 regressive movements, Number 56 required 891 fixations of which 241 were regressive. The reading time of Number 27 was 1752 twenty-fifths of a second as compared with 6504 for Number 56. Neither was the difference due to the quality of the reading performance, for Number 27 apparently read with a higher degree of comprehension than Number 56. Their comprehension scores (based on the laboratory exercises) were 37 and 22, respectively.

TABLE XXXII
SUMMARY OF EYE-MOVEMENT DATA RELATING TO GROUP I

Student Number	Number of Fixations*	Number of Regressive Movements*	Reading Time in 25ths of a Second*	Compr. Score**
27	266	32	1752	37
7	310	28	1841	35
14	322	49	2258	45
26	347	22	2153	39
1	347	50	2161	44
19	350	56	2302	36
22	350	59	2456	31
25	358	38	2223	37
29	363	47	2801	34
34	365	56	2768	39
10	370	67	2621	35
28	381	7	2536	27
49	382	46	2724	33
62	383	25	2863	31
16	384	73	2353	44
15	384	79	2829	40
8	391	57	2868	33
17	406	64	2511	38
33	411	66	3363	41
5	434	66	2798	37
4	440	65	3135	40
41	446	44	3690	44
6	446	60	3775	36
46	454	41	3369	20
42	461	26	3125	31
Mean	382.0	48.8	2691	36.3

* For reading sixty lines, ten lines in each of six paragraphs.

** Based on the laboratory exercises.

Number 36, who had fewer fixations than any other student in Group II, made 54 fixations - or nearly one per line - more than Number 42, the last student in Group I. Only two students in Group I made more than 70 regressive movements, whereas in Group II only two students made fewer than 70 regressive movements. Similarly, only two students in Group I required more than 3500 twenty-fifths of a second for reading sixty lines, and only one student in Group II required fewer than 3500. The range in performance is striking: 266 to 891 fixations, 7 to 325

TABLE XXXIII

SUMMARY OF EYE-MOVEMENT DATA RELATING TO GROUP II

Student Number	Number of Fixations*	Number of Regressive Movements*	Reading Time in 25ths of a Second*	Compr. Score**
56	891	241	6504	22
51	823	141	8266	39
53	804	267	5815	33
60	797	251	5916	36
12	790	173	6099	17
63	786	325	6501	16
32	749	177	6693	37
54	705	192	5333	32
31	697	177	5998	34
58	673	195	4765	24
38	670	141	5720	36
47	645	145	5259	20
50	642	96	4424	22
3	626	136	4279	31
55	620	117	5060	24
52	616	123	4641	30
57	589	102	4085	39
48	563	94	4673	26
59	561	176	4691	30
61	553	74	3813	33
39	551	67	3375	32
37	534	63	4554	28
30	523	88	4764	25
40	521	75	4094	20
36	515	119	4229	31
Mean	657.8	150.2	5182	28.7

* For reading sixty lines, ten lines in each of six paragraphs.

** Based on the laboratory exercises.

regressive movements, and 1752 to 8266 twenty-fifths of a second in reading time. The means also demonstrate significant differences between Groups I and II: (1) fixations, 392.0 as compared with 657.8; (2) regressive movements, 48.8 and 150.2; (3) duration of fixations, 2691 and 5182 twenty-fifths of a second.

Relation between mechanics of reading and comprehension.- In general, the faulty mechanics in reading reflected by the data in Table XXXIII tend to be associated with inadequate comprehension. Likewise, good reading habits tend to be associated with satisfactory comprehension. The relationship, however, is not perfect. Students 38, 32, 60, and 51 showed as much understanding of the paragraphs as the average reader in Group I, although each made from 11 to 14 fixations per line. Moreover, a few students with fairly good mechanical habits in reading showed very little appreciation for the thought content of what they had read. Note especially the low comprehension scores of Number 46 and Number 28 in Group I. The general relationship between good mechanical habits and adequate comprehension is indicated by the difference between the means in comprehension: for Group I, 36.3; for Group II, 28.7. That is, the students in Group I made the better records with respect to the mechanics of reading, yet read more thoroughly than the students in Group II.

Report of results, paragraph by paragraph.- Tables XXXIV to XXXIX inclusive show the performance of each student paragraph by paragraph. Tables XXXIV and XXXV deal with fixations per line; Tables XXXVI and XXXVII with regressive movements per line; Tables XXXVIII and XXXIX with the average duration of fixations in each paragraph. These data show how reading habits are modified whenever the reader's purpose is changed or when the subject matter is modified as it was in this experiment. Although the reader's performance reflects the influence of these variables, the data in the tables give an impression of uniformity rather than diversity. Each table is expressive of sameness, not only with respect to individuals, but also with respect to the two groups of students. The significance of the fluctuations, on the one hand, and of the stability of reading habits, on the other, will become clearer as the data are analyzed.

Since the study is primarily one in group contrasts, the analysis will be presented largely in terms of group performance. However, a few individual cases will be considered before turning to group comparisons.

Individual performance.- In so far as mechanical habits are concerned, Students 27 and 7 excel Number 14. However, their comprehension scores on the exercises used in the laboratory were only 37 and 35 respectively, whereas Number

TABLE XXXIV
AVERAGE NUMBER OF FIXATIONS PER LINE, GROUP I

Student Number	Paragraph Number						All
	II	III	IV	V	VI	VII	
27	4.6	4.8	4.4	3.6	3.9	5.3	4.4
7	5.0	5.1	5.2	6.1	4.3	5.3	5.2
14	5.0	6.3	5.1	5.4	4.7	5.7	5.4
26	5.2	6.0	5.5	6.4	5.8	5.8	5.8
1	6.4	6.0	5.3	6.4	5.2	5.4	5.8
19	5.5	6.0	5.3	6.9	5.5	5.8	5.8
22	4.7	5.4	5.8	6.8	5.9	6.4	5.8
25	5.1	7.3	5.3	7.3	4.9	5.9	6.0
29	4.3	10.4	4.9	6.5	5.1	5.1	6.0
34	5.6	6.3	6.1	6.8	5.5	6.2	6.1
10	4.8	9.1	5.6	5.6	5.6	6.3	6.2
28	6.4	7.1	5.9	5.9	6.0	6.8	6.3
49	5.1	7.8	6.1	6.3	5.9	7.0	6.4
62	6.2	7.3	6.4	6.4	6.2	5.8	6.4
16	5.6	7.7	5.8	6.0	6.7	6.6	6.4
15	5.1	7.3	6.4	6.9	6.1	6.6	6.4
8	5.3	6.5	6.2	7.1	7.1	6.9	6.5
17	6.9	8.6	6.8	6.2	5.8	6.3	6.8
33	6.9	7.0	8.0	6.9	6.3	6.0	6.8
5	6.3	7.8	7.2	7.0	7.6	7.5	7.2
4	6.8	7.6	7.3	6.6	7.7	8.0	7.3
41	7.0	8.7	7.0	7.8	7.2	6.9	7.4
6	6.4	7.2	7.4	6.7	7.5	9.4	7.4
46	6.4	8.7	7.9	7.5	7.0	7.9	7.6
42	7.7	8.7	8.3	7.0	7.5	6.9	7.7
Mean	5.8	7.2	6.2	6.5	6.0	6.5	6.4

14 made a perfect comprehension score. Therefore, Number 14 may be taken as the best reader in Group I - third in the group on the basis of mechanical habits and excelling all others in interpreting the paragraphs. Number 56 represents the opposite extreme. This student made more fixations per line than any other in Group II and, with the exception of Numbers 53, 60, and 63, made more regressive movements per line. Moreover, only four students made lower comprehension scores. The contrast between Students 14 and 56 is so extreme that their individual performance

TABLE XXXV
AVERAGE NUMBER OF FIXATIONS PER LINE, GROUP II

Student Number	Paragraph Number						All
	II	III	IV	V	VI	VII	
56	11.6	16.6	16.9	14.4	12.7	16.9	14.8
51	13.1	13.0	13.8	16.8	11.5	14.1	13.7
53	8.7	14.7	17.7	14.2	11.5	13.6	13.4
60	10.8	15.0	13.1	14.2	10.9	15.7	13.3
12	11.0	13.5	15.5	14.0	12.8	12.2	13.2
63	10.4	14.3	15.1	13.5	12.3	13.0	13.1
32	12.3	13.6	11.8	11.8	12.9	12.5	12.5
54	9.6	12.4	11.1	10.8	12.3	14.3	11.7
31	10.7	11.3	13.5	10.5	12.0	11.7	11.6
58	9.7	13.5	13.3	10.2	9.9	10.7	11.2
38	10.4	12.2	13.5	11.4	10.2	9.3	11.2
47	10.0	12.1	12.3	10.2	9.6	10.3	10.7
50	10.7	11.0	10.6	11.5	9.7	10.7	10.7
3	8.9	11.7	10.9	10.1	9.4	11.6	10.4
55	8.8	12.3	10.5	10.6	9.6	10.2	10.3
52	10.1	10.4	10.6	10.4	9.9	10.2	10.3
57	9.4	10.5	10.3	8.7	10.3	9.7	9.8
48	8.3	10.6	10.1	9.7	7.7	9.9	9.4
59	7.9	9.0	9.8	9.6	9.0	10.8	9.3
61	8.1	9.6	9.4	9.7	8.5	10.0	9.2
39	9.5	9.7	9.1	8.3	9.7	8.8	9.2
37	9.4	8.7	8.0	8.5	10.2	8.6	8.9
30	8.4	8.7	8.9	7.8	8.8	9.7	8.7
40	7.3	9.4	9.1	8.8	8.8	8.7	8.7
36	7.8	10.1	7.2	8.3	7.6	10.5	8.6
Mean	9.7	11.8	11.7	11.0	10.3	11.3	11.0

records are presented in detail in Tables XL and XLI. Number 56 made 2.8 times as many fixations as Number 14. Number 56 read only 4 of the 60 lines with fewer than 10 fixations per line. Number 14 did not make 10 fixations in any line and in only 6 of the 60 lines as many as 8 fixations. The contrast with respect to regressive movements is even greater. Number 56 made approximately five times as many regressive movements as Number 14 - from 1 to 10 in every line. Number 14

TABLE XXXVI
AVERAGE NUMBER OF REGRESSIVE MOVEMENTS PER LINE, GROUP I

Student Number	Paragraph Number						All
	II	III	IV	V	VI	VII	
27	.7	.3	.7	.3	.4	.8	.5
7	.5	.5	.5	.7	0.0	.6	.5
14	.7	1.3	.7	.9	.4	.9	.8
26	0.0	.3	.3	.8	.4	.4	.4
1	1.1	.9	.6	1.1	.6	.7	.8
19	.5	.9	.9	1.5	.7	1.1	.9
22	.4	.8	1.1	1.3	1.0	1.3	1.0
25	.2	1.2	.4	1.3	.2	.5	.6
29	.1	1.9	.4	1.2	.6	.5	.8
34	.7	.7	1.5	1.3	.6	.8	.9
10	.5	2.0	.7	1.4	.9	1.2	1.1
28	.2	0.0	0.0	0.0	.2	.3	.1
49	.4	1.2	.6	1.2	.2	1.0	.8
62	.4	.5	.7	.5	.3	.1	.4
16	.9	1.7	1.3	.8	1.4	1.2	1.2
15	.6	1.4	1.5	1.6	1.2	1.6	1.3
8	.4	.8	.7	1.4	1.2	1.2	1.0
17	1.2	1.3	1.3	1.1	.5	1.0	1.1
33	.9	1.0	1.4	1.2	1.0	1.1	1.1
5	.8	.9	1.2	1.3	1.2	1.2	1.1
4	.6	1.5	1.0	.6	1.4	1.4	1.1
41	.5	1.4	.4	1.1	.5	.5	.7
6	.8	.9	1.0	.5	.9	1.9	1.0
46	0.0	1.0	1.0	.9	.6	.6	.7
42	.6	.5	.8	.4	.3	0.0	.4
Mean	.5	1.0	.8	1.0	.7	.9	.8

read 24 lines without any regressive movements and made not more than 3 in any line. The demands made on the muscles which control the eye are, in the one case, very heavy as compared with the other. Fatigue would undoubtedly result more quickly. Finally, the duration of the fixation pauses in the case of Number 56 was slightly greater than that of Student 14. Moreover, with all these differences, Number 14 made a comprehension score of 45 and Number 56 a score of only 22.

TABLE XXXVII
AVERAGE NUMBER OF REGRESSIVE MOVEMENTS PER LINE, GROUP II

Student Number	Paragraph Number						All
	II	III	IV	V	VI	VII	
56	2.5	4.5	4.9	4.0	3.2	5.0	4.0
51	2.0	1.9	2.7	3.9	1.4	2.2	2.4
53	2.3	4.6	6.5	4.9	3.6	4.8	4.5
60	2.8	4.8	4.2	5.3	3.1	4.9	4.2
12	2.4	2.7	3.7	3.6	2.2	2.7	2.9
63	3.6	6.1	6.9	5.2	4.9	5.8	5.4
32	2.6	3.9	2.6	2.9	2.8	2.9	3.0
54	2.0	3.3	3.1	2.3	3.7	4.8	3.2
31	2.5	3.1	3.3	2.4	3.0	3.4	3.0
58	2.6	4.3	4.2	2.8	2.6	3.0	3.3
38	1.7	3.1	3.7	3.1	1.5	1.0	2.4
47	2.2	2.7	3.3	1.8	2.2	2.3	2.4
50	1.7	1.6	1.7	2.4	.7	1.5	1.6
3	2.0	2.4	3.1	2.1	1.7	2.3	2.3
55	.9	2.6	2.7	2.4	1.3	1.8	2.0
52	1.6	2.1	1.9	2.5	2.0	2.2	2.1
57	1.1	2.1	1.8	1.8	2.1	1.3	1.7
48	1.1	1.7	2.2	1.9	.9	1.6	1.6
59	2.2	2.9	3.2	2.9	2.8	3.6	2.9
61	.9	1.3	1.4	1.2	1.1	1.5	1.2
39	.9	1.2	.9	1.1	1.2	1.4	1.1
37	.8	.9	.9	.9	1.6	1.2	1.1
30	1.3	1.6	1.6	1.2	1.4	1.7	1.5
40	1.0	1.5	1.5	1.2	1.1	1.2	1.3
36	1.7	2.5	1.2	2.0	1.7	2.8	2.0
Mean	1.9	2.8	2.9	2.6	2.2	2.7	2.5

The significance of these facts is apparent. It is wholly absurd to expect comparable achievement from these two students when that achievement depends on their ability to use reading as a tool for study. The one can proceed approximately three times as fast as the other and at the same time accomplish a great deal more.

Individual differences in ability stand out clearly when Plate I is contrasted with Plate II. The eye-movements of the first five students in Group I

TABLE XXXVIII
AVERAGE DURATION OF FIXATIONS IN TWENTY-FIFTHS OF A SECOND,
GROUP I

Student Number	Paragraph Number						All
	II	III	IV	V	VI	VII	
27	6.6	7.5	6.4	6.4	5.9	6.5	6.6
7	5.8	6.1	6.0	5.8	6.2	5.8	6.0
14	6.6	7.3	6.6	7.3	7.1	7.0	7.0
26	5.8	6.4	5.9	6.2	6.2	6.6	6.2
1	6.0	6.1	6.7	6.4	6.1	6.1	6.2
19	6.7	7.2	6.1	6.2	6.7	6.7	6.6
22	6.9	6.8	7.2	7.0	7.1	7.0	7.0
25	5.6	6.8	5.6	6.1	6.7	6.4	6.2
29	7.3	7.9	7.3	8.0	7.6	7.8	7.7
34	7.5	8.7	6.8	6.8	7.0	8.7	7.6
10	6.6	7.9	7.0	6.6	6.8	6.9	7.1
28	6.8	6.7	6.6	6.7	6.5	6.6	6.7
49	6.3	7.3	7.4	7.2	7.4	7.0	7.1
62	7.5	7.6	7.4	7.3	7.3	7.7	7.5
16	5.8	6.9	6.2	6.3	5.8	5.6	6.1
15	6.4	8.3	7.1	7.4	7.4	7.5	7.4
8	7.4	7.7	7.3	7.2	7.2	7.2	7.3
17	6.0	6.6	5.8	6.1	6.3	6.2	6.2
33	7.9	8.8	8.7	7.9	7.6	8.1	8.2
5	5.8	7.2	6.4	6.5	6.4	6.2	6.4
4	7.0	7.9	6.7	7.6	6.8	6.7	7.1
41	7.9	8.1	7.9	8.3	8.3	9.1	8.3
6	10.0	7.9	7.5	7.5	9.2	8.7	8.5
46	7.2	7.8	7.3	7.7	7.4	7.1	7.4
42	6.6	7.3	6.4	7.0	6.7	6.6	6.8
Mean	6.8	7.4	6.8	6.9	6.9	7.0	7.0

on lines 8 and 9 from Paragraph II have been plotted in the usual manner in Plate I.¹ These specimens are typical of the performance of the best readers. Comparable records appear in Plate II for the first five readers in Group II. The difference between Plates I and II is evident.

Student 27² read Paragraph V with fewer fixations per line than any other

TABLE XXXIX
AVERAGE DURATION OF FIXATIONS IN TWENTY-FIFTHS OF A SECOND,
GROUP II

Student Number	Paragraph Number						All
	II	III	IV	V	VI	VII	
56	7.3	7.6	7.2	7.0	7.4	7.4	7.3
51	9.4	9.5	9.6	10.7	9.6	11.3	10.0
53	6.7	7.4	7.2	7.9	6.7	7.2	7.2
60	7.5	6.9	8.1	7.4	7.0	7.5	7.4
12	7.9	7.4	7.7	8.2	7.6	7.5	7.7
63	7.5	8.8	8.3	7.9	8.3	8.7	8.3
32	8.6	9.0	8.6	9.5	8.9	9.0	8.9
54	7.8	7.2	7.1	8.1	7.8	7.6	7.6
31	8.1	8.6	8.7	8.9	8.6	8.8	8.6
58	6.5	7.2	7.2	7.4	7.1	7.1	7.1
38	8.3	8.7	8.3	8.2	9.2	8.6	8.5
47	8.2	8.4	7.5	8.1	8.2	8.6	8.2
50	6.2	7.1	6.0	7.2	7.1	7.7	6.9
3	6.3	7.1	6.9	6.6	6.7	7.3	6.8
55	9.4	8.2	7.9	8.4	7.4	7.8	8.2
52	6.7	7.6	8.3	7.6	7.7	7.3	7.5
57	7.0	7.6	6.7	6.9	6.6	6.7	6.9
48	8.0	8.3	8.3	8.3	7.7	9.0	8.3
59	7.3	8.6	8.0	8.8	8.5	8.9	8.4
61	6.6	7.1	6.6	7.3	7.0	6.9	6.9
39	5.9	6.6	6.2	6.2	5.7	6.2	6.1
37	9.2	8.6	8.8	8.0	8.6	7.8	8.5
30	8.5	9.6	8.8	10.1	8.7	9.0	9.1
40	7.2	7.6	8.1	8.6	7.6	8.1	7.9
36	8.5	8.6	8.7	7.8	7.7	7.9	8.2
Mean	7.6	8.0	7.8	8.0	7.7	8.0	7.9

¹Each vertical line indicates where a fixation was made. The numerals above the line indicate the serial order in which the fixations occurred. Those below indicate the number of twenty-fifths of a second that the eye remained fixed at each point.

²See Table XXXIV, p. 90

TABLE XL

INDIVIDUAL PERFORMANCE RECORD OF STUDENT 14

Para-graph No.	Line Number *										Av. No. Fix.	Av. No. R.-Mov.	25ths of a Sec.	Av. Duration
	4	5	6	7	8	9	10	11	12	13				
II	3	7 ₂	6 ₂	6 ₁	5	6 ₁	4	4	5 ₁	4	5.0	.7	331	6.6
III	7 ₁	6 ₁	4	7 ₂	7 ₂	4	8 ₃	8 ₂	8 ₂	4	6.3	1.3	462	7.3
IV	5 ₁	4	6 ₁	5 ₁	4	5 ₁	4	5	4	9 ₃	5.1	.7	338	6.6
V	5 ₁	4	9 ₂	6 ₁	6 ₁	6 ₁	4	5 ₁	4 ₁	5 ₁	5.4	.9	396	7.3
VI	6 ₁	4	4	5 ₁	4	4	4	4	6 ₁	6 ₁	4.7	.4	331	7.1
VII	6 ₁	5 ₁	4	6	4	7 ₂	8 ₂	6 ₁	5 ₁	6 ₁	5.7	.9	400	7.0
Total **											322	49	2258	

*The number in each cell refers to the number of fixations made in reading that line. The subscripts indicate the number of regressive movements.

**Decimals omitted.

TABLE XLI

INDIVIDUAL PERFORMANCE RECORD OF STUDENT 56

Para-graph No.	Line Number *										Av. No. Fix.	Av. No. R.-Mov.	25ths of a Sec.	Av. Duration
	4	5	6	7	8	9	10	11	12	13				
II	12 ₃	12 ₃	9 ₁	9 ₁	14 ₄	13 ₄	11 ₁	14 ₃	11 ₂	11 ₃	11.6	2.5	849	7.3
III	22 ₇	15 ₄	14 ₃	17 ₅	12 ₂	18 ₇	25 ₇	16 ₄	15 ₄	12 ₂	16.6	4.5	1254	7.6
IV	11 ₃	22 ₇	24 ₈	16 ₅	27 ₁₀	10 ₂	14 ₄	14 ₃	14 ₃	17 ₄	16.9	4.9	1214	7.2
V	20 ₆	14 ₃	8 ₁	13 ₃	13 ₃	16 ₄	20 ₇	12 ₄	13 ₅	15 ₄	14.4	4.0	1008	7.0
VI	19 ₈	12 ₂	8 ₁	12 ₄	14 ₃	12 ₂	14 ₄	11 ₂	11 ₃	14 ₃	12.7	3.2	935	7.4
VII	14 ₄	20 ₇	22 ₈	17 ₅	20 ₆	11 ₂	18 ₅	14 ₃	17 ₅	16 ₅	16.9	5.0	1244	7.4
Total **											891	241	6504	

*The number in each cell refers to the number of fixations made in reading that line. The subscripts indicate the number of regressive movements.

**Decimals omitted.

Plate I

severance. The pathway to success is not a royal road. It does
not lead down hill, nor even across the level prairie. It is a steep

severance. The pathway to success is not a royal road. It does
not lead down hill, nor even across the level prairie. It is a steep

severance. The pathway to success is not a royal road. It does
not lead down hill, nor even across the level prairie. It is a steep

severance. The pathway to success is not a royal road. It does
not lead down hill, nor even across the level prairie. It is a steep

severance. The pathway to success is not a royal road. It does
not lead down hill, nor even across the level prairie. It is a steep

Silent Reading by Five Superior Readers

The performance of superior readers is illustrated by samples taken from the records of Students 27, 7, 14, 26, and 1 (the first five students in Group I) for lines 8 and 9 of Paragraph II.

Plate II

severance. The pathway to success is not a royal road. It does

not lead down hill, nor even across the level prairie. It is a steep

severance. The pathway to success is not a royal road. It does

not lead down hill, nor even across the level prairie. It is a steep

severance. The pathway to success is not a royal road. It does

not lead down hill, nor even across the level prairie. It is a steep

severance. The pathway to success is not a royal road. It does

not lead down hill, nor even across the level prairie. It is a steep

severance. The pathway to success is not a royal road. It does

not lead down hill, nor even across the level prairie. It is a steep

Silent Reading by Five Poor Readers

The performance of inefficient readers is illustrated by samples taken from the records of Students 56, 51, 53, 60, and 12 (the first five students in Group II) for lines 8 and 9 of Paragraph II.

student on any paragraph. Plate III is illustrative of his performance. He made only 36 fixations in reading the 10 lines, or 3.6 fixations per line.

Number 53¹ made the greatest number of fixations, 177 pauses in reading ten lines of Paragraph IV. Plate IV illustrates that type of performance. Number 28² made fewer regressive movements than any other student. Three of the paragraphs were read without any regressive movements. His comprehension score was not high. Perhaps this student should have made more regressive movements than he did. Plate V is plotted from Paragraph IV, one of those in which he made no regressive movements.

Plate VI presents a striking contrast to Plate V. It shows how many regressive movements were necessary in one instance. Number 63³ made 69 regressive movements in reading Paragraph IV, the largest number made by any student on any paragraph.

Students 25 and 16⁴ share honors with respect to short duration pauses. In each case the student had an average of 5.6 twenty-fifths of a second for one or more paragraphs. Plate VII is taken from the film of Number 16 on Paragraph VII.

The longest average duration of fixations occurred in the case of Student 51⁵ in reading Paragraph VII. Plate VIII may be contrasted in this connection with Plate VII.

The averages summarized.- The averages from Tables XXXIV to XXXIX are brought together in Table XLII in order that group comparisons may be more easily made. The averages denote a number of significant facts.

Number of fixations made by mature readers.- Superior readers average about six or seven fixations per line in reading material such as that used in this

¹See Table XXXV, p. 91

²See Table XXXVI, p. 92

³See Table XXXVII, p. 93

⁴See Table XXXVIII, p. 94

⁵See Table XXXIX, p. 95

Plate III

cause of *gravity* and that fire rose because of *levity*. Modern phys-
icists speak of four-dimensional space, of protons and electrons,
of fugacity and relativity, and of a host of other minutiae which
are utterly incomprehensible to the uninitiated. The abstractness
of science makes it arduous for the neophyte. When its terminol-

Silent Reading with Few Fixations per Line

Student 27 read Paragraph V with few fixations per line.
His record for lines 9 to 13 is shown.

Plate IV

it, then one compliments one's self on the exercise of such remark-
able judgment. As a matter of fact it was not judgment at all;
it was impulse, prejudice. Much of man's thinking goes on in
exactly this way. The sum total of one's habitual prejudices is his
viewpoint. When we say that another person is *right*, we mean

Silent Reading with Many Fixations per Line

Student 53 made a great many fixations in reading Paragraph IV
as illustrated by this record for lines 9 to 13 inclusive.

Plate V

it, then one compliments one's self on the exercise of such remarkable judgment. As a matter of fact it was not judgment at all; it was impulse, prejudice. Much of man's thinking goes on in exactly this way. The sum total of one's habitual prejudices is his viewpoint. When we say that another person is *right*, we mean

Silent Reading with Few Regressive Movements per Line

Student 28 read with few regressive movements. He made none in reading Paragraph IV. Lines 9 to 13 are included in this example.

Plate VI

it, then one compliments one's self on the exercise of such remarkable judgment. As a matter of fact it was not judgment at all; it was impulse, prejudice. Much of man's thinking goes on in exactly this way. The sum total of one's habitual prejudices is his viewpoint. When we say that another person is *right*, we mean

Silent Reading with Many Regressive Movements per Line

Student 63 made many regressive movements in reading Paragraph IV. Contrast this record with Plate V especially with respect to regressive movements.

Plate VII

activities. Its experiences have a warm intimacy and an in-
 separable concreteness which cannot be shared with any other
 person. It is only imperfectly and by analogy that one self ap-
 prehends the inner life of another. This fact of ownership and
 unitary experience occasionally comes to light in that marvelous

Silent Reading with Short Fixation Pauses

Student 16 made very short fixation pauses. His record for lines 9 to 13 of Paragraph VII is shown.

Plate VIII

activities. Its experiences have a warm intimacy and an in-
 separable concreteness which cannot be shared with any other
 person. It is only imperfectly and by analogy that one self ap-
 prehends the inner life of another. This fact of ownership and
 unitary experience occasionally comes to light in that marvelous

Silent Reading with Long Fixation Pauses

Student 51 made very long pauses. Contrast the numbers below the line in this record with those below the line in Plate VII.

experiment.¹ The poorer readers make ten to twelve fixations per line. If Group I is as good as Group II is poor, the averages for the two groups combined (the entire group of 50 students) should be rather typical of the average Chicago freshman. On that basis, these data indicate that the average Chicago freshman ordinarily makes eight to nine fixations per line.

TABLE XLII

A SUMMARY OF GROUP AVERAGES

Factors Considered	Group Number	Paragraph						All Paragraphs
		II	III	IV	V	VI	VII	
Average Number of Fixations	I	5.8	7.2	6.2	6.5	6.0	6.5	6.4
	II	9.7	11.8	11.7	11.0	10.3	11.3	11.0
	Both	7.8	9.5	9.0	8.8	8.2	8.9	8.7
Average Number of Regressive Movements	I	.5	1.0	.8	1.0	.7	.9	.8
	II	1.9	2.8	2.9	2.6	2.2	2.7	2.5
	Both	1.2	1.9	1.9	1.8	1.5	1.8	1.7
Average Duration of Fixations	I	6.8	7.4	6.8	6.9	6.9	7.0	7.0
	II	7.6	8.0	7.8	8.0	7.7	8.0	7.9
	Both	7.2	7.7	7.3	7.5	7.3	7.5	7.5

Plate IX presents samples selected from the records of Students 41 and 12 on Paragraph II as typical of the average performance of Groups I and II, respectively. Plate X from the record of Number 48 was selected as representative of the average performance of the two groups combined. In so far as the data are representative, this exhibit may be said to typify the reading habits of the average Chicago freshman.

Does the mature reader make extensive use of word grouping?— The foregoing estimate raises an important issue. It is not in agreement with a belief that is rather prevalent, namely, that the average mature person reads almost exclusively by word groups. Much of the literature on reading habits is based on the assumption that the mature reader makes only five or six short fixations per line, which

¹ The type of content, size of type, and length of line should be borne in mind in connection with these facts.

Plate IX

are willing to work and by those who never acknowledge defeat.

One of the finest contributions which your college work can make

to you is the habit of attacking your work vigorously and pursu-

are willing to work and by those who never acknowledge defeat.

One of the finest contributions which your college work can make

to you is the habit of attacking your work vigorously and pursu-

Silent Reading Typical of Groups I and II

The average performance for Group I is illustrated by the record of Student 41 for lines 11 to 13 of Paragraph II. (Part 1. above.)

The average performance for Group II is illustrated by the record of Student 12 for lines 11 to 13 of Paragraph II. (Part 2, above.)

Plate X

not lead down hill, nor even across the level prairie. It is a steep

and rugged pathway which can be traveled only by those who

are willing to work and by those who never acknowledge defeat.

One of the finest contributions which your college work can make

to you is the habit of attacking your work vigorously and pursu-

Silent Reading by a Student of Average Ability

The average performance for both groups combined is illustrated by the record of Student 48 for lines 9 to 13 of Paragraph II.

means, of course, that the reader must read by word groups. As a matter of fact, in this study only a few of the very best readers in Group I could be said to read in that fashion. The average for both groups, 8.7 fixations per line, is scarcely fewer than one fixation per word. Standards ordinarily set up as typical of the mature reader would apply only to the very best readers in Group I. The data suggest that the average mature reader is, to a considerable degree, a word-by-word reader.

Agreement between test scores and laboratory records.- Test results substantiate the fact that these data are typical of average performance. Allowing 8.7 fixations per line and 7.5 twenty-fifths of a second per fixation, the reader would require approximately 65 twenty-fifths of a second or two and three-fifths seconds for reading one line. In 100 seconds he could read about 38 or 39 lines or about 375 to 400 words. The average reader in the freshman class reads approximately at that rate. If it is assumed that six fixations per line of the same duration is the normal performance, then the average reader's rate should be 550, at least 1.5 words per second faster than the average obtained by means of the rate tests.

Other standards indicated by the data in Table XLII.- Similar differences exist between the observed data concerning regressive movements and duration of fixations and the standards which frequently appear in the literature on reading. That is, the average college freshman in this group made more regressive movements and each fixation pause was somewhat longer than might have been anticipated in view of results which have appeared in previous investigations. The averages in Table XLII indicate that good readers make one regressive movement in nearly every line, that average readers make two in many lines, and that poor readers make from two to three regressive movements in every line. The average duration of fixations ranges from 7.0 to 8.0 twenty-fifths of a second, the average for both groups being 7.5 twenty-fifths or three-tenths of a second.¹

¹See footnote 2, p. 84.

Meaning of the observed facts.- These data do not imply in any way that present standards are desirable. Perhaps the average reader in future classes can be, and will be, taught to read in the same way that the superior reader now proceeds. Word grouping should certainly be encouraged to the point where each individual has reached optimum efficiency. Word-by-word reading is none the less deplorable because it happens to be used more widely than is often assumed. The persistence of word-by-word reading for a decade after the possibilities of improvement in silent reading have been conclusively demonstrated is not surprising. On the other hand, it is well to recognize the fact that at present the average college student reads words much more frequently than is often supposed when he is reading selections written in language appropriate for the mature reader.

Probable causes of the difference in results reported.- It is the last-named factor which probably accounts for the observed discrepancy between current statements and the data obtained in this study. Norms established on simple paragraphs have been widely accepted as typical of general reading ability and given interpretations not intended by those who made the studies. For example, it is one thing to show how growth occurs from one grade level to another in the ability to read a given selection, as Buswell¹ has done. It is another to discover the typical performance of students at each grade level on material appropriate for the degree of maturity represented. Consequently, the fact that students make approximately six fixations per line of about six twenty-fifths of a second duration when reading simple material should not be interpreted to mean that such reading habits are typical. Judd and Buswell's² later investigation gives ample evidence of that fact. Nevertheless, the idea persists that the growth standards which Buswell and other investigators have announced are applicable generally. The following quotations contrast the materials used in Buswell's earlier study with those used in this investigation and show why the averages do not agree:

¹Op. cit.

²Op. cit.

- (1) "One night Peter went to bed early. It was not dark. The bright moon shone in at the window. Peter could see everything in the room. All at once he heard a noise.....etc."
- (2) "The average man, through his habitual prejudices, can be played upon by a skillful speaker or writer almost as readily as a musical instrument can be played upon by a skillful musician.....etc."

A second reason why these results differ from Buswell's is the fact that in his earlier study three and one-half inch lines were used instead of four-inch lines as in this experiment.

Characteristic differences between Groups I and II.- The averages in Table XLIII show that the factor which differentiates most clearly between Groups I and II is the number of fixations per line. The average number of fixations for Group II is four to five per line more than that for Group I.

The next set of averages in the table, those relating to regressive movements, seem to suggest that a considerable portion of the difference is due to the necessity for re-reading among students in Group II. Even the good reader makes regressive movements - one in nearly every line - but the inefficient reader makes approximately three times as many as the superior reader. In fact, inaccuracies in interpretation which make regressive movements necessary account for more than one-third of the observed difference between Groups I and II.¹ That regressive movements do not account for all the difference indicates that the perception span of the poor reader is somewhat narrower.

There is a slight but consistent advantage with respect to duration of fixations in favor of Group I. This is convincing evidence that the poor reader does not make a greater number of fixations because he makes them faster as certain theorists have advocated. The compensation idea finds no support in these data. Not only does the superior reader make fewer fixations per line, but he also pauses in shorter time intervals. The fact that the difference in average

¹Group I averaged 6.4 fixations per line; Group II, 11.0 - a difference of 4.6 fixations. The corresponding difference with respect to regressive movements was 1.7, which is more than one-third of 4.6.

duration of fixations for Groups I and II is small, implies that speed of recognition as an inherent characteristic is not as important a factor in reading as difficulty in interpretation. For, if the reader were merely slow in recognizing words, it would seem likely that there would be less difference between Groups I and II with respect to number of fixations and regressive movements and a greater difference with respect to duration of fixations than was true in this experiment. If the foregoing interpretation is correct, reading difficulties among college students are due primarily to the mental processes involved in thought-getting and only secondarily to slowness in word recognition. Emphasis in correcting them should therefore be on thought-interpretation rather than on speed in word recognition.

Differences in performance on various paragraphs.- The eye-movement records obtained with Paragraph II provide standards with which the data from other paragraphs may be compared. Paragraph II was simple, and when reading it the student was relatively free to adopt his own habitual style of reading. The means show that mechanical habits were at their best in the case of this paragraph. With each of the other paragraphs, certain conditioning factors were operating. The resulting differences, suggested by the means presented in Table XLII, are pointed out below.

With Paragraph III the reader's purpose was changed. Instead of "reading merely to understand" he was making "a vigorous effort to get all the important ideas." As the work of Judd and Buswell¹ would lead one to expect, there were more fixations per line, more regressive movements per line, and the fixation pauses were longer when the student was asked to read more carefully. Both groups were affected by the change in purpose. Generally speaking, the reader tended to make about two more fixations per line, to have one more regressive movement per line, and to pause about one-fiftieth of a second longer for each fixation. As the reader's purpose is intensified, his eye-movements are slowed down, the perception span is narrowed, and progress along the line is less regular. Incidentally, the change

¹Op. cit.

introduced by this factor was greater than that resulting from any other variable.

In reading Paragraph IV the student was again asked to read "merely to understand," but the injunction to read rapidly was less convincing after the student had just attempted (often rather unsuccessfully) to answer the questions on Paragraph III. Particularly was that true of Group II, whose average performance on Paragraph IV was only a trifle better than that on Paragraph III. The data show that the reading habits of Group II were less flexible than those of Group I; they were unable to adapt themselves readily to new reading situations.

The performance of the two groups on Paragraphs IV and V was something of a paradox. Although in reading Paragraph V, Group I made more fixations per line and more regressive movements than in Paragraph IV, Group II made fewer fixations and fewer regressive movements. In other words, the group that should have been most affected by difficult words seemed to be least affected by them, probably because of the carry-over effect of careful reading to Paragraph IV. In both groups the averages for Paragraph V were greater than those for Paragraph II, indicating that difficult words do have a perceptible effect on the reading habits. But in view of the facts just cited, it appears that vocabulary difficulties are less influential than the purpose of the reader.

Paragraph VI, which differed from the remaining paragraphs with respect to sentence structure, produced less variation than any of the others. A few students whose language difficulties were probably due to the fact that a foreign language was spoken in the home had difficulty with Paragraph VI. In general, however, sentence structure seemed to present little difficulty to the readers in either group. The averages indicate greater fluency for Paragraph VI than for Paragraph IV.

Paragraph VII, which dealt with more abstract ideas than any of the others, produced considerable variation in reading habits. The eye-movements in this instance were approximately equal in quality to those for Paragraph V, in which vocabulary difficulties were introduced.

Three of the four variables introduced showed considerable influence over

the performance of the reader. In the order of their importance the variables were: (1) purpose of the reader, (2) abstractness of the ideas, (3) difficulty of the vocabulary, and (4) complexity of the sentence structure.

Certain intercorrelations.- Correlations based on the eye-movement records of fifty students may be regarded as suggestive, at least. Table XLIII summarizes a number of coefficients which have been computed on the basis of entries in Tables XXXIV to XXXIX. Number of fixations per line was quite reliable as a measure of reading habits. The lowest correlation (.829) was found between the number of fixations in the criterion paragraph and in that in which the reader's purpose was modified. The uniform size of the coefficients indicates their dependability.

TABLE XLIII
CORRELATION COEFFICIENTS BETWEEN CERTAIN EYE-MOVEMENT MEASURES

Factors Correlated	Number Cases	r ± P. E.
Fixations Paragraph II with Paragraph III	50	.829 ± .03
" " II " " IV	50	.863 ± .02
" " II " " V	50	.876 ± .02
" " II " " VI	50	.889 ± .02
" " II " " VII	50	.860 ± .02
Regressive Movements Par. II with Par. III	50	.867 ± .02
Duration Paragraph II with Paragraph III	50	.735 ± .04
No. of Fixations (Av.) with No. of Regr. Mov. (Av.)	50	.880 ± .02
" " " " " Duration (Av.)	50	.435 ± .08
" " Regr. Mov. " " " "	50	.345 ± .08
Reading Time (Tables XXXII and XXXIII) with Rate (Table XXXI).	46	.801 ± .04

Since the correlation was lowest between Paragraphs II and III with respect to number of fixations per line, these paragraphs were also used in estimating the probable reliability of number of regressive movements and duration of fixations. This was done in order that the coefficient might represent the probable minimum. The correlation between the number of regressive movements on Paragraphs II and III was of the same general magnitude as the coefficients relating to the number of

fixations per line. The correlation for duration of fixations on the two paragraphs was somewhat lower, yet high enough to indicate a rather high degree of constancy.

The next three coefficients in the table portray the interrelationship between the three factors. The coefficient .88, although spuriously high, bears out the fact previously stated, that a large number of fixations per line tends to be accompanied by a large number of regressive movements per line. The last two coefficients show a positive but low relationship between the duration of pauses and the number of fixations or number of regressive movements. Although, in general, good readers make briefer pauses than poor readers, the difference is less pronounced than in the case of number of fixations or number of regressive movements. Good readers sometimes make longer pauses than poor readers. That fact is shown not only by the low coefficient just cited but also by the following illustrations from Tables XXXVIII and XXXIX. Student 39 in Group II averaged only 5.7 twenty-fifths of a second per fixation on Paragraph VI, while Student 6 in Group I averaged 9.2 twenty-fifths on the same paragraph. Five students in Group II averaged less than 7.0 twenty-fifths of a second on the six paragraphs, that being the average for Group I. Moreover, three students in Group I averaged more than 7.9 twenty-fifths of a second on the six paragraphs, the average of Group II. That the difference between the means is only .9 is, within itself, indicative of the fact that duration of fixations is relatively unimportant in distinguishing the good from the poor readers in these groups.

The last coefficient in Table XLIII shows the relation between reading time on 60 lines measured by the photographic technique and the rate score obtained from the entrance test. Since four students had not taken the entrance test, the coefficient is based on only 46 cases. Nevertheless, a coefficient of .80 between two independent measures of rate, the second of which was obtained after an interval of more than three months, signifies that reading rate tends to remain stable. A rate score is, therefore, a significant type of measurement. Moreover, since the laboratory measure is known to be entirely objective, this coefficient is evidence

of the validity of the rate score obtained with the Test of Achievement in Silent Reading.

Summary

Data pertaining to the reading tests developed in this investigation indicate that:

1. The typical Chicago freshman makes the following test scores: rate, 4.00; vocabulary, 30; comprehension, 48; and composite score, 118.
2. Two-thirds of the students make rate scores ranging from 3.00 to 5.00; vocabulary scores from 20 to 40; comprehension scores from 40 to 56; and composite scores from 100 to 140.
3. As tentative minimum standards, the following scores may be accepted: rate, 3.00; vocabulary, 20; comprehension, 40; and composite score, 100.
4. Tremendous individual differences in reading ability exist. Reading rates range from less than two to more than seven words per second (2.00 to 7.00), vocabulary scores from 8 to 48, comprehension scores from 20 to 66, and composite scores from 60 to 170.
5. Intelligence and general reading ability are closely related (coefficient .702). Intelligence correlates most highly with vocabulary (.724), next with comprehension (.569), and least with reading rate (.312).
6. Vocabulary and comprehension are closely related factors (zero-order coefficient, .607; partial correlation coefficient with intelligence constant, .343).
7. The correlation between rate and comprehension is positive but low (all coefficients lower than .30); likewise, that between rate and vocabulary (coefficient .38).
8. Reading ability is definitely related to academic achievement.¹

¹The correlation between reading test scores and grade points was almost as high as that between intelligence and grade points. Moreover, a significant positive correlation was found even when intelligence was held constant. Clear evidence of the relationship was also found in the reading deficiencies of students who withdrew from the University because of failure and in the unsatisfactory achievement of students who made unsatisfactory scores on the reading tests.

9. Approximately 10 per cent of Chicago freshmen are seriously handicapped in the three major phases of reading, rate, vocabulary, and comprehension, and as many more show marked deficiencies with respect to one or two of these factors.

The eye-movement records of twenty-five superior readers and twenty-five poor readers suggest that:

1. For superior readers the following averages prevail: fixations per line, 6.4; duration of fixations, 7.0 twenty-fifths of a second; regressive movements, .8 per line or about one regressive movement in eight fixations.

2. Comparable averages for poor readers are: fixations per line, 11.0; duration of fixations, 7.9 twenty-fifths of a second; regressive movements, 2.5 per line - every fourth or fifth fixation.

3. The typical Chicago freshman averages eight or nine fixations per line, each lasting three-tenths of a second (7.5 twenty-fifths), and of the eight or nine eye-movements, one or two are regressive.

4. Individual differences in the fundamental reading habits of college students are pronounced. The average number of fixations per line ranges from 4.4 to 14.8; regressive movements, from .1 to 5.4; and duration of fixations, from 6.0 to 10.0 twenty-fifths of a second.

5. Faulty mechanical habits in reading are frequently, but not invariably, accompanied by poor comprehension.

6. There is a distinct relationship between the quality of the fundamental reading habits and intelligence.

7. In the order of their importance, four variables which cause readers to make more fixations per line and more regressive movements are: (1) purpose of the reader, (2) abstractness of the material, (3) difficulty of the vocabulary, and (4) complexity of sentence structure.

8. Good and poor readers differ more with respect to number of fixations per line than with respect to either number of regressive movements or duration of fixations.

9. Number of fixations per line based on ten lines of material is a very

reliable index of reading ability (reliability coefficients ranging from .83 to .89).

10. There is a positive relation between number of fixations per line and number of regressive movements (coefficient .88); also a low but positive correlation between the number and duration of fixations (.435) and between duration of fixations and number of regressive movements (.345).

The reading test scores and eye-movement records were in close agreement with respect to: (1) the identification of poor readers, (2) the amount of difference in reading ability which exists at the college level, (3) the typical reading habits of Chicago freshmen, (4) the relation between rate and comprehension in reading, (5) the relation between intelligence and general reading ability, and (6) the reliability of a rate score. The data presented in this chapter provide insight into the reading habits of college students and reaffirm the need for remedial training among them.

CHAPTER IV

AN EXPERIMENT IN GROUP INSTRUCTION

Relation of the Remedial Work to Other Phases of the Investigation

Earlier phases of this study were concerned with the need for remedial instruction; the final step, with the development and evaluation of a particular type of instruction. By developing and evaluating a new method of measuring silent reading achievement and by using this and other instruments to analyze the reading abilities and disabilities of Chicago freshmen, the need for remedial reading at the University of Chicago was clearly revealed and the scope of the problem determined in so far as that institution was concerned. The question remained, What should the University do to correct them? In view of the great variety of available procedures, obviously a complete answer to this broad question could not be found in this investigation. However, a first step could be taken; one type of remedial instruction could be employed as an experiment, its effectiveness determined, and its merits and defects noted. The present chapter describes an experiment directed toward this purpose.

Before turning to a description of the remedial work, however, its relation to other phases of the whole investigation will be further indicated by reviewing the order in which the activities of this study were carried out.

Sequence of activities. - A general survey of reading ability in the 1930-31¹ Chicago freshman class was made in October 1930. The Test of Achievement in Silent Reading, the first part of the battery developed in this study, was administered to all entering freshmen. This provided a measure of reading rate and a partial measure of comprehension, but not an adequate measure of reading ability in all phases. Since the other two parts of the new test battery were not yet available,

¹

This survey has been mentioned in previous sections of this report.

further information about reading achievement was derived from various sources. Other test scores and records of academic achievement were consulted for supplementary evidence relating to the reading abilities and disabilities of the students.¹

In the light of the facts discovered, students in the 1930-31 freshman class were classified in four groups: (1) those whose records showed rather clearly that serious reading difficulties were present, (2) "doubtful cases," or those whose records suggested, but did not prove, that reading difficulties were present, (3) average readers, and (4) superior readers. In November, follow-up tests were given to the 255 students in Groups 1, 2, and 4: 61 from the first group, 108 from the second, and 86 from the fourth.² In this work the Minnesota Reading Examination³ and the newly completed Test of Reading Vocabulary were employed.

Data pertaining to the number of students with serious reading deficiencies who were identified in this survey were presented in Chapter III. Among

¹ Among the entrance tests were: the American Council Psychological Examination and the Sones-Harry High School Achievement Test. Percentile rank on the former indicated in a general way the student's ability in academic fields - in which reading is an important element. In addition to the total score on the latter, the scores on the following parts of the test proved helpful: (1) Part A, entitled "English Usage," (2) Part G, "Comprehension," (3) Part B, "General Vocabulary," and (4) Part L, "Technical Vocabulary of Language." The student's rank in his high-school graduating class, the general average of all his high-school grades, his high-school average in the social studies, and his high-school average in English were other factors considered. The average scholarship mark for each student was obtained from calculations made under the direction of Karl J. Holzinger for use in the Preliminary Report on the Application of the Sones-Harry and Other Tests to University Freshmen. Unpublished report, November, 1930. The grades assigned under different marking systems were "brought to a common basis by projecting all grades on a scale beginning at 75 as passing and running up to 100. Letter grades were projected in a similar way under the assumption of a normal curve." Grades in the social studies and in English were considered because reading demands are relatively heavy in these fields.

² Since this investigation is not based on the performance of the "average reader," the third group was not included in the follow-up program.

³ For complete citation see p. 6.

them were the students who participated in the remedial project with which the present chapter deals. Further details with reference to their selection will be given presently. Both the superior readers and the poor readers whose eye-movement records were presented in Chapter III were likewise selected on the basis of data obtained in this survey.

Following the work just described, the battery of reading tests was completed, the laboratory work reported in Chapter III was done, a study was made of the reading difficulties of each student who expected to enroll for the proposed remedial instruction, and plans for the remedial work were developed, including the necessary administrative arrangements. All these activities were integral parts of the study as a whole; all had to be completed before the remedial work could begin. However, these matters made it necessary to postpone the remedial instruction until the spring quarter, that is, until approximately the first of April, 1931. The significance of this fact will become apparent in the following discussion of the experiment.

Plans for the Remedial Work

Type of instruction - Two possible types of remedial work were considered in connection with this investigation: individual instruction, and group instruction. Each method has certain merits that the other cannot claim. In individual instruction maximum attention can be devoted to individual differences in interests and capacities. Materials and methods can be selected with the needs of a particular individual in mind. During the period of instruction the remedial worker can give undivided attention to the learner, directing his activities to better advantage and noting his errors more readily. On the other hand, group work keeps the individual from feeling that he is being singled out. Comparison of individual achievement with that of the group provides the learner with a wholesome incentive to further effort. And finally, group instruction makes it possible for one remedial worker to instruct a larger number of students.

Consideration of the advantages and disadvantages of the two methods and of the possible contribution of this experiment led to the selection of group instruction as the type to be used in this study. In the first place, the individual or case-study method of giving remedial instruction in reading has been used extensively and many techniques appropriate for use with individuals have been carefully evaluated.¹ In the second place, group instruction was more appropriate in view of the large number of students whose reading disabilities were known to constitute a real handicap. It was recognized that, in so far as the 1930-31 class was concerned, one worker could not give remedial training to all who needed it by individual instruction nor reach all who needed the training with a single remedial class. However, it seemed apparent that if the University expected to bring eventually to all handicapped readers a program of successful remedial instruction, some kind of group training would have to be employed with a majority, if not all, such students. Consequently, an experiment in group instruction seemed to constitute the best departure in attacking the problem of remedial instruction in reading at the University of Chicago. It would determine the feasibility of group instruction and, by revealing the merits and shortcomings of a particular method of instruction, could be used as a guide for further experimentation.

Specific purposes of the remedial work- The remedial instruction was undertaken with three objectives in view: (1) to develop and evaluate one method of group instruction, (2) to aid the students participating in the project in building up better reading habits, and (3) to determine the practical utility of the reading tests described in Chapter II in diagnosis and in the measurement of progress.

Selection of a specific group method. - Remedial workers dealing with groups of students have, in previous studies, adopted one or the other of two widely different plans. One of these may be illustrated by the procedures of

¹ See especially, William S. Gray, Remedial Cases in Reading: Their Diagnosis and Treatment. Supplementary Educational Monographs, Number 22. Chicago: Department of Education, The University of Chicago, 1922.

Mrs. Pressey.¹ The emphasis in her work was on the development of specific skills one after another. A manual of special exercises and drill materials was used. The other type of procedure is illustrated by Parr's experiment.² In his study emphasis was placed on thought-getting rather than on the mechanics of reading, and the materials of a regular college course were used as practice materials for improving reading habits.

The latter method was selected for use in this study. Its selection was based on the hypothesis that, at the college level, the mental processes involved in comprehension rather than the mechanics of reading are of primary importance.

Therefore, plans were developed for a program of instruction in which emphasis would be placed on comprehension and organization in reading. Such a course would help to show whether the attention of college students should be focused on eye-movement habits and the development of specific mechanical skills, or whether attention to thought-getting is sufficient. In this course, neither speed drills, flash-card exercises designed to lengthen the perception span, nor any other formal device to improve a specific mechanical habit in reading would be appropriate, however useful such procedures are in certain instances. If remedial instruction of this type were to prove satisfactory or even feasible, the value of the project was apparent. Remedial classes could be organized in connection with regular college courses and the reading materials of the latter could be used in the development of better reading habits. The time devoted to reading would not be an extra burden on the student, but an integral part of his regular work. Moreover, if reading habits could be improved with instruction of this type, the work could readily be broadened to include other forms of directed study and perhaps other essential forms of guidance. On the other hand, if the method is not a good one, its shortcomings should be pointed out.

In view of these facts, inquiry was made concerning a regular college course which seemed well adapted for a cooperative project; a course in which the

¹ Op. cit.

² Op. cit.

assigned reading materials might conceivably be used as practice materials in a class looking to the improvement of silent reading habits.

Selection of a college course with which to cooperate. - Early in the winter quarter, sixty of the poorest readers in the 1930-31 freshman class were selected on the basis of their test scores on the entrance and follow-up tests mentioned above.¹ Each student's tentative program for the spring quarter was then recorded in order to discover the class in which the greatest number of inefficient readers would be enrolled. This class proved to be History 141, a course entitled "A Survey of Modern Civilization." Since the history department generously offered its cooperation, plans were made for a remedial reading class in conjunction with this survey course.

For several reasons History 141 seemed well adapted to the proposed work in remedial reading. In the first place, it required extensive reading. Each student was asked to procure and use a rental set of nine volumes.² Second, the materials used in the course varied greatly in content and style, ranging all the way from straightforward narration to difficult philosophical treatises. Third, the purpose

1

During the autumn quarter, or at the close of it, several of the poorest readers withdrew from the University. The sixty students mentioned here include those remaining of the sixty-five handicapped readers mentioned on page 78 and a few other students whose scores were low in several instances.

2

The following books were in the rental set:

- Charles A. Beard, Whither Mankind. New York: Longmans, Green & Company, 1929. pp. 408.
Barrett H. Clark, Great Biographies of the World. New York: Robert M. McBride and Company, 1928. pp. 1408.
G. Lowes Dickinson, The International Anarchy, 1904-1914. New York: The Century Company, 1926. pp. 506.
Frederick C. Dietz, The Industrial Revolution. New York: Henry Holt & Company, 1927. pp. 112.
Helen Gardner, Art Through the Ages. New York: Harcourt, Brace & Company, 1926. pp. 506.
L. T. Hobhouse, Liberalism. New York: Henry Holt & Company, 1911. pp. 254.
Shailer Mathews, The French Revolution, 1789-1815. New York: Longmans, Green & Company, 1927. pp. 466.
John Herman Randall, The Making of the Modern Mind. Boston: Houghton Mifflin Company, 1926. pp. 654.
Ferdinand Schevill, A History of Europe from the Reformation to Our Own Day. New York: Harcourt, Brace & Company, 1925. pp. 728.

of the course seemed favorable to the cultivation of better reading habits. History 141 was one of three survey courses recently developed to give freshmen a broad overview of world history. Success in it did not depend on the mastery of a narrow or technical field. Of course, training in the social sciences would give certain students an advantage over others who lacked such training, but there was no reason why any student should fail to master the essentials of the course merely because he lacked the proper background. Students were expected to acquire insights, understandings, and appreciations rather than to memorize a series of facts. Finally, the content of the course seemed likely to be especially interesting. Being the third in the series, History 141 dealt with the origin and characteristics of modern civilization.

Selection of cases for the remedial class. - The sixty poor readers were notified that a class in remedial reading was being organized in conjunction with History 141. A direct recommendation was made to those who were planning to enroll in the history course that they enroll in the reading class also. The remainder of them were invited to participate in the experiment if they could conveniently rearrange their programs to take History 141. Sixteen students volunteered for the work, fifteen of the sixty who were invited to enroll and one student who sought admission to the class on his own initiative.

Since all students in the remedial reading class also participated in the laboratory work described in Chapter III, they will be identified throughout this chapter by the same student numbers assigned in connection with the laboratory work and used in Tables XXXI to XXXIX.

Limits fixed by the experimental group. - The experimental group was smaller than the class originally planned. The work was organized in the hope that twenty-five to thirty-five of the sixty poorest readers would participate. Early replies to the invitation mentioned above justified that expectation, but several students withdrew before time for the work to begin. Some of the poorest readers withdrew from the University. Some who were interested in the proposed training found it impossible to adjust their programs of study to

include History 141. Other students who needed the training and might have taken it did not volunteer to do so.

It was unfortunate that the remedial program could reach only one-fourth of the students whose reading difficulties were known to constitute a serious handicap. It would also have been better from the standpoint of the project itself if a larger experimental group had been available. Conclusions based on a larger number of cases could have been drawn with somewhat greater certainty than were warranted from the data obtained. However, since a larger class could not be organized at that time in the year, the work was continued with the sixteen students who could participate. This situation imposed a definite limit on the remedial project but did not make its fundamental objectives impossible of attainment.

A control group. - A small but carefully selected control group was used in this investigation. Ten of the students who participated in the laboratory study discussed in Chapter III but not in the remedial class, were paired with ten students in the remedial class on the basis of the number of fixations made in reading sixty lines before the eye-movement camera. It should be recalled that: (1) the laboratory measurements were made before the remedial work began, (2) eye-movement records constitute the most objective measure of reading ability yet devised, and (3) number of fixations was found to be the most reliable index of reading ability yielded by the eye-movement records.

The data in Table XLIV show how the students in the control group compared with the ten students selected from the remedial class with respect to: (1) total number of fixations made in reading sixty lines, (2) total number of regressive movements, and (3) percentile rank on the American Council Psychological Examination. The averages show a slight advantage in favor of the control group with respect to both reading ability and intelligence.

This control group falls short of an ideal one in certain respects. Sixteen cases would have been better than ten. The use of various criteria in addition to eye-movement records might have resulted in greater accuracy if each member of the control group could have been chosen from several whose eye-movement records were

comparable. The use of more than one control group also has much to commend it. None of these refinements were possible in this investigation. Under the circumstances, the ten-student group, chosen on the basis of eye-movement records, seemed to afford the most accurate control that could be established.

TABLE XLIV

EYE-MOVEMENT AND INTELLIGENCE RECORDS OF TEN STUDENTS IN THE REMEDIAL READING CLASS COMPARED WITH THOSE OF TEN OTHER STUDENTS

Student Number.		Number of Fixations in Reading 60 Lines*		Number of Regressive Movements in 60 Lines*		Percentile Rank on Intelligence*	
Group A (Remedial)	Group B (Control)	A	B	A	B	A	B
51	53	923	804	141	267	35	02
60	12	797	790	251	173	28	06
63	32	786	749	325	177	28	57
54	31	705	697	192	177	57	—
58	38	673	670	195	141	—	35
55	3	620	626	117	136	35	86
37	30	534	523	63	88	10	43
36	40	515	521	119	75	17	06
62	15	383	384	25	79	28	—
28	16	381	384	7	73	22	94
Mean		621.7	614.8	143.5	138.6	29.0	41.1

* Data from Tables XXXI, XXXII, and XXXIII, pp. 85, 87, and 88.

Further plans. - After selecting the college course with which the work in remedial reading was to be done and securing the cooperation of the students involved, the following problems were attacked: (1) the diagnosis of the reading difficulties of the sixteen students, (2) the formulation of a statement of

appropriate general and specific objectives for the work of the class, (3) the development, in so far as possible, of specific procedures to be used in class work, (4) an evaluation of the materials used in the history course, and (5) the selection of a time and place for the sessions of the remedial reading class and the completion of all other administrative arrangements. The discussion now turns to the first of the problems mentioned above.

The Diagnosis of Reading Difficulties

Ability of the group as a whole. - Data presented in Table XLV reveal a number of significant facts with reference to the remedial class. The first three columns merely identify the student as Number 50, Number 55, and so on,¹ and show that the class was made up of ten young men and six young women, all of whom were approximately eighteen years old. Column 4 shows the percentile rank of each student on the American Council Psychological Examination. Half of the students ranked lower than the twenty-fifth percentile, and only one student ranked higher than the fiftieth percentile.

Columns 5 and 6 show the students' scores on the Test of Achievement in Silent Reading, given in October, 1930. Six of the rate scores are higher than 3.00, the standard of minimum achievement recommended in Chapter III, but when the "Cr" scores are taken into account, it is apparent that five of these do not represent rate of efficient reading. The "Cr" scores of Numbers 47, 37, 61, 62, and 28 were not satisfactory. Consequently, it is probable either that these students attempted to read more rapidly than usual, or that they were in the habit of reading superficially. Column 7 gives the quartile rank of each student on the comprehension part of the Minnesota Reading Examination and Column 8 his

¹ Since these numbers correspond to the numbers used in Tables XXXI to XXXIX, the detailed eye-movement record of a particular student can be found readily. For example, Student Number 50 (the first student listed in Table XLV) is the thirteenth student listed in Tables XXXIII, XXXV, XXXVII, and XXXIX.

rank on the vocabulary part of the same test. Eight students were below the fourth or lowest quartile with respect to comprehension and thirteen were below the average of the freshman class as a whole. With respect to vocabulary, thirteen students ranked lower than the fourth quartile and only two students exceeded the class average.

Columns 9, 10, and 11 in Table XLV describe, in a general way, the eye-movement habits of the students in the remedial class. Nine of the sixteen students averaged more than ten fixations per line in reading the sixty lines

TABLE XLV
DATA CONCERNING STUDENTS IN THE REMEDIAL READING CLASS

Student Number	Sex	Age	Intelligence P. R.	Reading (Entrance Test)		Quartile Rank Minn. Test		Eye-Movement Record (Averages)			Rank on Assoc. Tests ^a	
				Rate	Cr.	Compr.	Vocab.	Fix.	Regr.	Dur.	Controlled	Free
1	2	3	4	5	6	7	8	9	10	11	12	13
50	boy	18	1	2.57	5	4	4	10.7	1.6	6.9	15	13
55	boy	18	35	2.57	5	3	4	10.3	2.0	8.2	13	15
59	boy	19	4	-	-	4	4	9.3	2.9	8.4	16	14
63	boy	17	28	2.58	4	4	4	13.1	5.4	8.3	14	12
52	girl	18	35	2.73	4	3	4	10.3	2.1	7.5	3.5	8.5
54	girl	17	57	2.69	6	4	4	11.7	3.2	7.6	2	3
48	boy	18	6	3.73	6	4	4	9.4	1.6	8.3	11	6
47	boy	18	17	3.32	5	4	4	10.7	2.4	8.2	7	2
37	girl	18	10	3.82	4	4	4	8.9	1.1	8.5	6	11
61	girl	18	17	4.08	4	4	3	9.2	1.2	6.9	1	4
60	girl	17	28	2.23	4	2	4	13.3	4.2	7.4	12	16
36	girl	17	17	2.41	7	1	4	8.6	2.0	8.2	10	5
51	boy	18	35	2.08	2	2	4	13.7	2.4	10.0	3.5	10
58	boy	18	-	-	-	3	2	11.2	3.3	7.1	8	8.5
62	boy	19	28	5.25	2	3	4	6.4	.4	7.5	9	7
28	boy	19	22	3.47	5	3	2	6.3	.1	6.7	5	1
Average		18	23	3.11	4.5			10.2	2.2	7.8		
Standard ^b			50	3.83	7.5			8.7	1.7	7.5		

^a A description of these tests and a discussion of the relation between association time and reading ability are reported by Arthur E. Traxler, op. cit., pp. 80-96.

^b Based on averages obtained with other groups.

of material used in the laboratory study. Only three students made fewer fixations per line than the average for the fifty students photographed. Ten students averaged more than two regressive movements per line, and for half the group the average duration of fixations was greater than eight twenty-fifths of a second. Such records indicate clearly that, for some reason, these students did not proceed regularly and easily in silent reading. Only two students, Numbers 62 and 28, exhibited good mechanical habits when reading before the eye-movement camera. The two final columns in the table show how the students in the class compared with one another in the matter of association time, both controlled association and free association. Although no standards were available for use with these tests, the rankings show which pupils were relatively slow in making mental associations. Traxler¹ has pointed out the importance of this factor in connection with reading rate.

Combining the evidence summarized above with other test scores and other records which were available, it was possible to determine with reasonable assurance: (1) the characteristic reading difficulties of each student, (2) the apparent cause or causes of the difficulties, and (3) phases of reading to which the student should devote special attention. A diagnosis of the difficulties of each student is presented in the paragraphs which follow. As a matter of convenience in discussion, the cases have been classified arbitrarily on the basis of some one element of similarity into four groups.

Students who were deficient in all phases of reading. - The test scores and eye-movement records of Students 50, 55, 59, 63, 52, and 54 revealed the need of all-round improvement in reading. In so far as the apparent symptoms of difficulty were concerned, the cases appeared to have much in common. Note their uniformly poor achievement as shown by the data in Table XLVI. Further analysis of the cases, however, revealed marked differences in the apparent causes of

¹Op.cit., pp. 80-96.

difficulty. Statements of these primary causes, in so far as they could be determined, and of the remedial measures outlined for each student follow.

TABLE XLVI
ACHIEVEMENT RECORDS OF SIX STUDENTS LOW IN ALL PHASES
OF READING

Student Number	Entrance Test in reading		Quartile Rank Minn. Test ^a		Eye-Movement Records ^a		
	Rate	Cr	Compr.	Vocab.	Fix.	Regr.	Dur.
50	2.57	5	4	4	10.7	1.6	6.9
55	2.57	5	3	4	10.3	2.0	8.2
59	---	---	4	4	9.3	2.9	8.4
63	2.58	4	4	4	13.1	5.4	8.3
52	2.73	4	3	4	10.3	2.1	7.5
54	2.69	6	4	4	11.7	3.2	7.6

^a Data from Table XLV, p. 125.

Number 50 was one of the most handicapped readers in the entire Chicago freshman class. Only one student in the remedial group made lower vocabulary and comprehension scores on the Minnesota Reading Examination. Analysis of his vocabulary troubles showed that they did not pertain to word recognition, if recognition is interpreted as the ability to pronounce words. In oral reading Number 50 demonstrated moderate ability to pronounce words. Note also that his eye-movement record in Table XLVI shows relatively few regressive movements. His difficulty was that words often carried no meaning or an incorrect meaning. He recognized words, but not word meanings. The primary cause of his reading difficulties, of which the vocabulary difficulty mentioned above was but one, seemed to be inherent. On the American Council intelligence test his percentile rank was 1, placing him at the very bottom of his class with respect to native capacity. His general average in high school was 90.4, but he was rated by the principal as belonging in the "middle third" of a graduating class of 730 students. His scores on the Sones-Harry general achievement test were also very low. He ranked thirteenth with respect to rate of

free association and fifteenth on controlled association. Of course, none of these data indicated that the student was subnormal. But, neither did his reading scores indicate a type of reading that would be inferior except in a selected group. Consequently, lack of native capacity seemed to be the primary cause of the boy's poor reading habits.

Since obviously direct attack on the primary cause of the difficulty was impossible, plans for the remedial work with student Number 50 were made as follows: (1) to observe his work carefully for evidences of lip-movement, line following, or any other habit that might interfere with or retard his reading, (2) to assist him as much as possible in developing a larger reading vocabulary, (3) to encourage him to make as much use as possible of word grouping, but discourage him in attempting in a rate too rapid for adequate comprehension, and (4) to give the student practice in using various aids to interpretation.

Number 55 surpassed Number 50 very little in reading achievement, but his difficulties were of a different nature. It seemed reasonably apparent that the reading habits of this student were not developed to the limits set by his native capacity. His percentile rank on the intelligence test was 35, not a high rating, but equal or superior to nearly 200 of his classmates. His rank on both association tests was low: thirteenth on the controlled association test and fifteenth on free association. He had taken little interest in general reading in high school. He said that extra-curricular activities took so much of his time that he did not have much time to read. His favorite subject in high school was history. He had never been required to read widely in any high-school course. His inefficient habits of reading were apparently the result of exclusive use of a more or less superficial type of textbook study.

In view of the student's slow rate of association, little emphasis on rapid reading would be warranted. The problems of greatest importance in the remedial work seemed to be: (1) the development of a better attitude toward reading and the stimulation of wholesome interest in books, (2) guidance in interpreting various types of material, and (3) assistance in building up a

larger reading vocabulary.

Number 59 was especially handicapped with respect to vocabulary. His score on the Minnesota test was 10 out of a possible score of 100. Not only was he unfamiliar with the meaning of words, but also with their pronunciation. Vocabulary difficulties were indeed such a hindrance to this student that, at times, he seemed not to have formed the reading adaptation. Under the pressure of a test situation he would pass rapidly from one word to the next, even though the result was but a meaningless jargon. When he really attempted to read understandingly, his rate of reading was slow. His performance in the laboratory was erratic. Much of the time during the laboratory experiment he did not read understandingly, but rather in the manner described above. He averaged 9.3 fixations and 2.9 regressive movements per line. There was considerable irregularity in his performance from line to line. The necessity for constant attention to the recognition and meaning of words frequently caused him to ignore punctuation marks. This fact became apparent from his performance on Gray's Standardized Oral Reading Paragraphs. Moreover, in reading these paragraphs he made a larger number of mispronunciations and substitutions than an average sixth-grade pupil. The cause of his language disability was readily discovered. His parents were foreign born. The foreign language was spoken exclusively in the home. The student's only acquaintance with the English language had been acquired in school and community activities. He had succeeded in meeting the requirements of a large high school and, although his high-school record was not especially good, he had gained admission to the University of Chicago. His percentile rank on the intelligence test was only 4. His rank on the controlled association test was sixteenth and on the free association test fourteenth. Accompanying the environmental handicap mentioned above, these factors were by no means reassuring.

This student needed remedial work directed toward the development of a larger reading vocabulary. Other phases of language, such as sentence interpretation and the significance of punctuation marks, called for similar emphasis. The habit of "pretending to read" was to be discouraged. Most important of all, Number 59 needed to read extensively from materials far less difficult than those assigned

in college classes. Practice in oral and written expression was also badly needed, as his language handicap was distinctly noticeable even in his conversation.

Number 63 had a percentile rank of 28 on the American Council intelligence test. His rank on both association tests was relatively low. Neither of these factors, however, seemed to account for his reading disabilities. His small vocabulary was evidently a hindrance to rapid reading and to satisfactory comprehension, but this was not the primary cause. Examination of his high-school record showed that he had taken only the required courses in English and in the social studies and had made his lowest grades in these departments. His best work was done in science and in mathematics. He stated that he seldom read for pleasure. He was interested in sports, especially tennis, but not in books. Further inquiry disclosed the fact that oral reading was virtually the only type of reading instruction which he had ever had. He could get the meaning of a difficult passage or paragraph better by reading it aloud than by reading it silently. Apparently his present faulty habits were but the continuation of improper procedures developed in the elementary school and reinforced by a growing dislike for reading as time went by.

In view of these facts, the first problem in remediation seemed to be the creation of a better attitude toward reading. This boy also needed a great deal of practice in the silent reading of interesting and relatively easy material. He required guidance in approaching each reading situation with a definite purpose in mind. He needed instruction on how to adapt his reading to different purposes and to different types of material.

Number 52 was another student whose reading difficulties could be traced to over-emphasis on oral reading in her early training. This possibility was suggested by the lack of flexibility in reading habits demonstrated in the laboratory work. Whether reading "as rapidly as possible" or "carefully and thoughtfully" and regardless of the material at hand, this student made approximately ten fixations and two regressive movements per line. According to her report, oral reading was the only type in which she had ever received formal instruction. There was one important difference between this case and that of Number 63 described

above. This student liked to read and, in spite of her methodical habits, had read considerably more than was actually required of her. Her best work in high school was in English and the social studies. Her percentile rank on the intelligence test was 35. On both association tests her rank was above the average of the remedial group. Consequently, the special objectives of the remedial work with this student were: (1) guidance in adapting her method of reading to different purposes and different types of material, and (2) stimulation of greater interest in vocabulary development.

Number 54 was a student whose poor reading achievement was somewhat surprising in view of her excellent high-school record and relatively high scores on the Sones-Harry general achievement test. Her high-school average was 94.4 and she ranked twelfth in a graduating class of 192 students. Her percentile rank on the intelligence test was 57. Her ranks on the association tests were 2 and 3, respectively. Retests showed that, given plenty of time, the student could interpret passages in a very satisfactory manner. Her low comprehension scores on earlier tests were due in part to lack of time. Two significant causal factors were discovered. The first was the habit of dawdling. This was the student's own diagnosis of the case and it was apparently correct. She expressed considerable surprise when her low reading test scores were called to her attention. No one had ever told her that she was a poor reader and the idea had never occurred to her that she could develop better reading habits. She said that in high school she read as little as possible, acquiring information in other ways. By being alert in class, taking notes, and entering freely into discussions, she found little real need for reading. Consequently she had plenty of time to do what little reading seemed advisable, and she read accordingly. The second contributing cause was a language handicap. The student's parents used a foreign language exclusively. The children of the family had learned English in the public schools and had adopted it, even in the home, except when speaking to their parents. Hence, this student was by no means as unskilled in English expression as student Number 59, referred to above. Nevertheless, this situation probably accounted in part for her limited vocabulary.

In the remedial work, special emphasis was needed with respect to: (1)

adapting the style of reading to different purposes, (2) using helps to interpretation such as headings, topic sentences, and summaries, (3) grouping words into thought units, (4) developing a larger reading vocabulary, and (5) the habit of reading more extensively.

Satisfactory rate; inadequate comprehension and vocabulary. - Four students, Numbers 48, 47, 37, and 61, made satisfactory rate scores on the Test of Achievement in Silent Reading but ranked low with respect to comprehension and vocabulary. The symptoms of difficulty apparent from the test scores on the survey tests and from the students' eye-movement records may be noted from the data in Table XLVII. All the "Cr" scores were low; and all the comprehension scores and three of the vocabulary scores based on the Minnesota test were below the fourth quartile.

TABLE XLVII
ACHIEVEMENT RECORDS OF FOUR STUDENTS HAVING LOW COMPREHENSION
AND VOCABULARY SCORES

Student Number	Entrance Test in Reading ^a		Quartile Rank Minn. Test ^a		Eye-Movement Records ^a		
	Rate	Cr	Compr.	Vocab.	Fix.	Regr.	Dur.
48	3.73	6	4	4	9.4	1.6	8.3
47	3.32	5	4	4	10.7	2.4	8.2
37	3.82	4	4	4	8.9	1.1	8.5
61	4.08	4	4	3	9.2	1.2	6.9

^aData from Table XLV, p. 125

These cases were similar in more respects than those in the group previously described. In every case, vocabulary difficulties were sufficiently pronounced to be regarded as a cause of poor comprehension. The fact that reading rate was satisfactory suggested that the vocabulary difficulties pertained to word meanings primarily, rather than to recognition in the narrowest sense. Eye-movement records superior to those of many other students in the remedial group and reasonably good

oral reading ability (demonstrated by these students in informal follow-up tests) confirmed this hypothesis. Intelligence was perhaps a minor factor, the percentile ranks being 6, 17, 10, and 17 respectively. Student 61 ranked high in the matter of association time; the others, about average among students in the remedial class. Number 37 had the best record of high-school achievement, an average grade of 92.1 and a rank in class of 42 among 397 students. No one of the four students manifested special interest in reading or reported having read extensively in connection with high-school courses. Analysis of their errors on the Test of Accurate Interpretation in Reading showed that most of them occurred on the parts dealing with "total meaning" and "organization" and relatively few on the part dealing with "fact questions." This fact suggested that the habit of reading without sufficient attention to the content was an important cause of difficulty in each case.

The phases of remedial instruction selected for special emphasis with these students were: (1) attention to word meanings, with a view to developing a larger reading vocabulary, (2) reading with a purpose clearly in mind, (3) the selection of important ideas, (4) the organization of materials, especially mental organization while reading, and (5) adapting the reading processes to the purpose and type of material at hand.

Slow readers whose comprehension was satisfactory. - The reading habits of four students, Numbers 60, 36, 51, and 58, were opposite in character to those of the students just described. Instead of normal rate and poor comprehension, the group under discussion combined satisfactory comprehension with a reading rate below an acceptable standard. The symptoms of their difficulty as summarized in Table XLVIII stand in marked contrast to those presented in Table XLVII. Numbers 60 and 51 proceeded more slowly than Numbers 36 and 58. Furthermore, their "Cr" scores were low on the Test of Achievement in Silent Reading even at the rate adopted. Their eye-movement records also betrayed faulty mechanical habits in reading. Numbers 60, 36, and 51 made low scores on the vocabulary test. In spite of such difficulties, these students were able to derive meaning from a selection. Number 36 surpassed all members of the remedial group on comprehension. Number 58

exhibited satisfactory knowledge of word meanings. Differences in the apparent causes of the difficulties involved merit a brief diagnosis of each of these cases.

TABLE XLVIII
ACHIEVEMENT RECORDS OF FOUR SLOW READERS HAVING
SATISFACTORY COMPREHENSION

Student Number	Entrance Test in Reading ^a		Quartile Rank Minn. Test ^a		Eye-Movement Records ^a		
	Rate	Cr	Compr.	Vocab.	Fix.	Regr.	Dur.
60	2.23	4	2	4	13.3	4.2	7.4
36	2.41	7	1	4	8.6	2.0	8.2
51	2.08 _b	2 _b	2	4	13.7	2.4	10.0
58	2.86	6	3	2	11.2	3.3	7.1

^a Data from Table XLV, p. 125

^b Number 58 took the Test of Achievement in Silent Reading as an individual test because he entered the University too late for the entrance tests. Since these scores were obtained under conditions not strictly comparable to those for other students, they are not reported in Table XLV.

Number 60 encountered difficulty in word recognition. This was undoubtedly one cause of her slow rate of reading. In the laboratory, she made more fixations and more regressive movements per line when reading the paragraph that contained difficult words than when reading other paragraphs. In oral reading, too, she had difficulty in pronouncing words. Another factor which may have influenced her reading rate was her association time. Her rank on both association tests was low: twelfth on the controlled, and sixteenth on the free association test. The primary cause of her difficulties, however, was probably a limited background of reading experience. This element was suggested by the student's low scores on the Sones-Harry general achievement test and supported by her own testimony with reference to reading interests and habits. She stated that she had always been, and still was, more interested in sports and social events than in books.

It seemed apparent that the success of the remedial work with this student

would depend in no small measure on the development of greater interest in and a better attitude toward reading. In addition, emphasis was needed on: (1) making an independent attack on new words, (2) attending to word meanings, (3) reading for a specific purpose, and (4) adapting the reading processes to the purpose at hand.

Number 36 was a slow reader largely through force of habit. Because of her excellent comprehension scores, the individual follow-up tests used with this student consisted of having her read as long selections as possible in given time intervals and report orally on their content. On these tests she read much more efficiently than on the entrance test. In the laboratory work, also, where she realized the importance of reading well, her eye-movement habits compared favorably with the average for the fifty students involved. The most probable explanation for this student's habit of working below capacity seemed to be a vision defect. She was quite nearsighted. Without glasses she was unable to read 11-point type at a distance greater than ten inches. At the time of this investigation, and apparently for a number of years back, the student wore glasses to compensate for this defect. However, her reading disabilities were probably due, at least in part, to faulty habits developed before the difficulty was corrected.

Remedial work was especially needed to develop the habit of working at a higher level of efficiency. The instruction would need to stress: (1) approaching each reading situation with a definite purpose in mind, (2) concentrating attention, (3) adapting the style of reading to different purposes and different types of material, (4) making greater use of interpretation helps such as headings, topic sentences, and summaries, and (5) developing a larger reading vocabulary.

Number 51 encountered serious vocabulary difficulties. These were apparent not only from his scores on vocabulary tests but from his oral reading and from the eye-movement records. The latter indicated two significant facts about his slow rate of reading. First of all, his perception span was narrow. He averaged 13.7 fixations per line, of which 2.4 were regressive movements. Second, his eye-movement pauses were unusually long in duration. Many of them were quite normal in length, but he frequently paused from ten to fifteen twenty-fifths of a second and

sometimes even longer. The average duration of his fixation pauses was 10.0 twenty-fifths of a second, one-third longer than the average found for fifty students. Analysis of the records revealed no single cause for the long fixation pauses. To some extent the longer pauses seemed to occur with the longer and more difficult words. It appeared, therefore, that this student employed a few long fixation pauses when confronted with a difficult word rather than several fixations and regressive movements. The association of long pauses with long but very common words such as "concreteness", "ownership", "experience", and "occasionally" suggested a second possibility, namely, that the student vocalized every word in reading. Since repeated observation revealed no tendency toward the use of lip-movement, this theory could be verified only by the student's self-analysis. He stated that, although he did not say each word with his lips, he was conscious of "hearing" each word as though he had pronounced it aloud. When asked to note any contractions of the throat muscles or tongue while reading silently, he reported that there was a marked tendency for him to vocalize. Although neither vocabulary difficulties nor the time required for vocalization seemed to account for all the long fixations - for some of them occurred on words such as "comes" and "of" - these factors were undoubtedly important.

According to the student's report of his early training, oral reading was the only type in which he had had any instruction. His silent reading was therefore merely a kind of silent oral-reading. The inflexibility of his reading habits gave further proof that such was the case. Except for the paragraph containing difficult words, where he made more fixations and regressive movements than usual, his eye-movements were quite similar for the various paragraphs.

To overcome the difficulties which have been described, remedial work was needed in breaking down the student's one-type procedure. Guidance was required in overcoming the tendency to say and hear each word. He needed especially to recognize the various purposes of silent reading and to learn how to adapt his reading to the purpose at hand. Special emphasis on working at maximum speed, on word grouping, on the use of aids to paragraph interpretation, and on vocabulary

development seemed to be appropriate also to the difficulties involved.

Number 58 had reading habits somewhat similar to those of Number 36 already described. He was quite interested in his test scores and eye-movement records and expressed great surprise at his relatively poor performance. The primary cause of his difficulty seemed to be that, in the absence of any information about his reading habits, he had been satisfied to work at a low level of efficiency. Number 58 was reared in a home where the parents used a foreign language exclusively. In view of this fact, his range of word knowledge seemed remarkable. His score on the Minnesota test was higher than the average for all Chicago freshmen and was next to the highest made by any student in the remedial group.

Since the remedial instruction needed by this student seemed to be almost identical with that already outlined for Number 36, specific details need not be given here.¹

Students with good mechanical habits. - Two students in the remedial group, Numbers 62 and 28, surpassed all the others with respect to the mechanics of reading.² Their eye-movement records were very similar. Number 62 averaged 6.4 fixations and .4 regressive movements per line, with fixation pauses of 7.5 twenty-fifths of a second. Number 28 averaged 6.3 fixations and .1 regressive movements per line, with fixations 6.7 twenty-fifths of a second in length. These records, standing alone, seem to indicate superior reading ability. Other evidence, however, revealed the fact that these students were superficial readers. Neither made satisfactory "Cr" scores on the Test of Achievement in Silent Reading, nor the Test of Accurate Interpretation used in follow-up work. Both showed more ability to answer fact questions than to organize materials or select and retain important ideas.

¹See p. 135. Special emphasis on the development of a larger reading vocabulary (the last item listed) was not required with this student.

²See Table XLV, p. 125 for the records of these students.

Number 62 had a limited reading vocabulary. One reason he frequently failed to acquire meaning was that he would skip unfamiliar words without pausing even to consider what they meant.

The primary cause of the difficulty in both cases seemed to be the students' attitude toward reading. Both students had developed the habit of hurrying through their reading assignments, satisfied with whatever ideas they chanced to glean. Number 28 was the student who sought admission to the remedial class on his own initiative. He was about to be suspended from the University because of failing grades and thought that reading difficulties might be responsible for his failures. Superficial reading was probably one element involved, but beyond that were interests and attitudes that affected all his work, whether reading was involved or not.

To the extent that the foregoing diagnosis is correct, remedial instruction was most needed to discourage superficial reading. Special emphasis was required on: (1) concentrating attention while reading, (2) reading for a purpose, (3) attending to word meanings, and (4) organizing, evaluating, and associating ideas. The development of a larger reading vocabulary by attending to word meanings was particularly important in the case of Number 62.

Uses made of case analyses.- The facts just presented pertaining to the nature of the reading difficulties of the sixteen students in the remedial group and the hypotheses suggesting the probable causes of them were used in three ways: (1) in developing a general plan for the work of the remedial class, (2) in adapting class procedures to individual needs, and (3) in acquainting students with their own abilities and disabilities and the probable causes thereof, in order that they might develop an effective plan of self-improvement. Since the last named use will not be mentioned elsewhere, a further word of explanation seems in place.

During the first week of the spring quarter, just before group instruction began, an individual conference was held with each student. In the interview, the student's difficulties and their apparent causes were discussed sympathetically but frankly. His scores were presented to him, not as isolated marks, but in relation

to each other. It was suggested that he undertake, on his own initiative, whatever steps toward self-improvement his program of studies would allow. He was told how he could probably improve his reading habits most.

Objectives for the Work in Remedial Reading

Two kinds of objectives were outlined for use in the remedial work: (1) certain general objectives which would pervade every reading period, and (2) certain specific objectives one or more of which would become the center of interest in each class exercise. Each objective was evaluated in terms of the following questions: (1) How many students in the remedial group need the type of training which this objective implies? (2) To what extent would failure to provide the training implied by this objective hinder development along other lines? and (3) Is this objective appropriate in view of other plans for the remedial work?

The specific objectives were so arranged that each division or group could be used as the basis for one unit of the remedial work. They were so organized that the class would proceed from the comprehension of larger thought units to that of smaller units; that is, the first group of objectives dealt with the comprehension of an entire passage or selection, the second with paragraph comprehension, the third with sentence interpretation, and the last with word meanings.

The objectives were stated in terms of student abilities. Pronouns in the second person were used in wording them in order to make their appeal to each student as direct and personal as possible. Examination of the following lists will show that major emphasis has been placed on types of remedial work indicated in the foregoing section of this chapter as most essential for these students, namely, the development of a favorable attitude toward reading, the habit of focusing attention on content, flexibility in reading habits, skill in interpreting different types of material in a variety of ways, and attention to word meanings.

GENERAL OBJECTIVES OF THE COURSE

- I. In order to awaken interest in reading and develop a wholesome attitude toward the work of the college to such an extent that you will work with maximum efficiency, you should be able to:
 1. Recognize your own handicaps.
 2. Discover a purpose for each reading situation.
 3. Control your attention and concentrate for a considerable period of time.
 4. Look upon reading as thought getting.
 5. Work at your maximum rate.
- II. In order to adapt the style of reading to your purpose, you should be able to:
 1. Skim merely for the run of the thought.
 2. Skim a passage in order to find specific words, phrases, or ideas.
 3. Read rapidly in order to acquire an author's point of view.
 4. Read rapidly in order to discover and adopt the author's organization.
 5. Make a rapid preliminary survey of material which you are going to read more carefully.
 6. Read carefully in order to gain new insight into a problem, or a new understanding of a principle.
 7. Read carefully to evaluate the material as you read. (Judging the truth of facts presented, the validity of the proof offered, the soundness of the logic involved, the aptness of the illustrations used, or the accuracy of the terminology employed.)
 8. Read for the purpose of finding the answer to a specific fact question.
 9. Select relevant facts from reference material.
 10. Review material previously studied.
 11. Skim a passage for unfamiliar and difficult words.
 12. Read silently, following a rapid oral reader.
 13. Read for the purpose of discovering new problems or formulating significant questions.
 14. Read critically in order to analyze the meaning of a difficult or unusual expression.
- III. In order to adapt your style of reading to the type of material at hand, and to develop habits of reading appropriate to the type of material as well as to the purpose you have in mind, you should be able to:
 1. Adapt your style of reading to familiar material.

2. Adapt your reading to unfamiliar material.
3. Adapt your reading to material which is simple in style as contrasted with that which is complex.
4. Adapt your reading to material which is concrete in nature.
5. Adapt your reading to material which is abstract.

SPECIFIC OBJECTIVES OF THE COURSE

- I. In order to comprehend an entire passage or selection, you should be able to:
 1. See the relation of a chapter or selection to a larger topic or unit of thought.
 2. Discover the central theme of a selection.
 3. Discover the author's organization.
 4. Make use of the standard devices and signs which indicate an author's organization.
 5. Make your own organization of a selection.
 6. See the relationship between the major and minor elements of thought in a selection.
 7. Make use of introductory and summary paragraphs.
 8. Detect whether adjacent paragraphs are on the same or different topics.
- II. In order to understand the meaning of a paragraph, you should be able to:
 1. See the relevance of a paragraph to a topic.
 2. Discover the main idea of a paragraph.
 3. Utilize the standard devices of paragraph organization, such as, headings and topic sentences.
 4. Recognize the contribution made by supporting details.
 5. See cause and effect relationships.
 6. Make use of transitional words, thus keeping pace with turns of thought.
 7. Visualize a scene or event depicted.
 8. Find the answer to a fact question.
 9. Organize the content of a paragraph.
- III. In order to understand the meaning of sentences, you should be able to:
 1. See the point of a sentence quickly.
 2. Select the principal parts of a sentence, the antecedent of a pronoun,

etc., where such analysis is necessary for accurate interpretation.

3. Read with sufficient care to be able to restate the idea in your own language.
 4. Interpret questions, problems, and directions.
 5. Phrase a sentence in thought units; i.e., to group the words into significant thought groups.
 6. Get meaning in spite of unfamiliar words.
 7. Utilize punctuation in interpreting meaning.
 8. Properly interpret transitional and connective words.
 9. See the relevance of a sentence to the paragraph of which it is a part.
 10. Recognize whether consecutive sentences are on the same or different ideas.
- IV. In order to understand the meaning of words, you should be able to:
1. Analyze difficult words as they are encountered.
 2. Utilize the context in discovering the meaning of unfamiliar words.
 3. Analyze words by the application of the usual meaning of the common prefixes, suffixes, and Latin stems.

Procedure Followed in Remedial Instruction

Class schedule. - The remedial reading class met one hour each school day for a period of eight weeks beginning April 6, 1931. The group did not meet during the first week of the quarter nor after May 29, 1931. Conferences with individual students, which were mentioned previously, were held at the scheduled class hour during the first week of the quarter. After May 29, the hour was used in re-photographing the eye-movements of the students.

The history class (History 141) met at 1:30. The remedial reading class met at 2:30 in the same room where, as Section J of the history class, they held their discussions on Mondays and Fridays.¹ A better time and place could scarcely

¹ More than ninety students - including the sixteen in the remedial reading class - were enrolled in History 141. On Tuesday, Wednesday, and Thursday of each week the entire class met in an auditorium for a class lecture. On Mondays and Fridays, small groups of ten to twenty students were scheduled to meet for discussion under the direction of the various instructors who assisted with the course. The sixteen students in the reading class constituted Section J of the history class under this plan of organization.

have been arranged. Two days a week the class simply remained in the same room after their regular work in history. Every day they came to the reading class with history more or less uppermost in their minds.

Test administered as the remedial work began. - The remedial work was launched by giving one form of the entire battery of reading tests described in Chapter II: Form 3 of the Test of Achievement in Silent Reading, Form A of the Test of Reading Vocabulary, and Form 300 of the Test of Accurate Interpretation in Reading. This step provided an important point of reference in evaluating the work to follow, for, after the earlier measurements were made, the students modified their reading habits in different ways. Some improved their habits considerably; others made little or no improvement. Since growth due to the remedial work could be distinguished from growth due to self-improvement during the autumn and winter quarters only by comparing achievement at the beginning of the term with that at the end of the term, these additional tests were essential.

The scores were also used by the remedial worker in reconsidering the difficulties and needs of students. Since, however, the case analyses based on previous measurements seemed essentially correct, the test scores need not be presented at this point. They will be given in the section which deals with the results of the remedial work along with the students' final scores on another form of the test.

Classroom procedures - Eighteen class periods were devoted to instruction and practice in interpreting entire chapters or sections. Ten class exercises were then given over to paragraph analysis - the second unit of the remedial work - five periods to sentence interpretation, and four to word analysis. The class time was allotted in this way because most of the students in the class seemed to need training in the interpretation of selections and paragraphs more than in the comprehension of isolated sentences and words.

To introduce the work in remedial reading, a copy of the statement of general and specific objectives (given on pp. 140-2) was given to each student. A discussion of the specific objectives followed. It was explained that, for several

days, the class work would consist of practice in interpreting whole sections of material or entire chapters; that materials of various kinds would be selected; that, each day, one or more of the specific objectives listed in the first division of that outline would become the objective of the class work; that the students would be asked to help select the specific objectives to be used from day to day.

The general objectives were discussed in somewhat greater detail. It was explained that one or more of these objectives would be referred to in practically every class period during the term. The students were asked to fasten both outlines of objectives in the syllabus used in History 141, to read them carefully, and to refer to them frequently in selecting appropriate objectives for reading done outside the remedial class. Three ideas were then developed as clearly and convincingly as possible: (1) that reading is thought getting, (2) that, since reading is thought getting, concentration of attention is necessary for effective reading, and (3) that the reader should always read with a purpose. The class periods which followed were devoted to reading and interpreting entire chapters or unified parts of chapters.

Each day the group was told which of the volumes from the rental set would be used on the following day. Materials were selected which bore directly on the topics under discussion in the history class. The remedial worker attended the lectures and the discussions of Section J in order to keep in constant touch with the work of the history class. The instructor in history recommended the materials which he regarded as particularly important. If at all appropriate in view of the abilities of the students and the objectives to be attained, selections were chosen from the materials thus recommended. From time to time, the students were asked to suggest selections which they would like to use in the following class period and the types of interpretation which the class should attempt. Extra copies of the book to be used on any given day were provided for students who, for any reason, came unprepared. Selections were chosen which could be read by the slower readers in twenty to thirty minutes. Additional material relating to the same general topic was selected for the use of the more rapid readers.

Before the class set to work, there was always a brief discussion of why the students were to read the selection before them. Then from their knowledge of the book of which it was a part, the title, and a hurried glance at the selection, the students attempted to predict something as to the nature of the material. Afterwards, in view of the purpose selected and the nature of the material at hand, they discussed the manner in which they should read the selection. Following the discussion the class set to work, reading silently the passage selected. The students were urged to focus attention exclusively on the task at hand, but speed of reading was not emphasized. No time records were kept. When the more rapid readers finished, they merely turned to the parallel material and continued to read. When practically all had finished the basic selection, the remedial worker would say, "Stop when you finish the paragraph you are now reading." The remainder of the period was then spent in interpreting the material in a manner appropriate to the objectives selected. Toward the beginning of the term, the students' comprehension of each selection was tested by oral discussions; for example, discussions of the relation between the selection read and the general topic under consideration in the history course, or perhaps, of the author's organization of the selection. Later, outlines and other types of written exercises were required. Throughout this unit of remedial work, considerable emphasis was placed on the organization of ideas. The final requirement was a two-page outline of the work covered thus far in the history course. Although the preparation of this outline was not an exercise in comprehending a single selection, the assignment seemed appropriate, for the work in history during the month constituted a complete unit. To construct a good outline, the student had to comprehend and properly evaluate many selections, integrating ideas from them with ideas acquired from the lectures and discussions.

1

The unit of work in history dealt with the role of certain revolutionary influences in ushering in the modern period of civilization. Significant intellectual developments, industrial and economic changes, and political revolutions were taken up one by one as forces which changed medieval civilization into that which we know today.

Certain provisions for individual differences in needs and abilities were made. Mention has already been made of the use of supplementary materials with the more rapid readers in the group. In addition, when a special reminder seemed appropriate in view of the difficulties of a particular student, the suggestion was given, sometimes privately, sometimes in group discussion. The following illustration may be cited. In one class exercise the group was preparing to read a selection entitled "The Industrial and Social Revolution of the Nineteenth Century,"¹ and to outline it, using the author's sectional and marginal headings as much as possible but modifying the author's organization to emphasize the points most significant in connection with that day's lecture in History 141. It was suggested definitely that students Number 48, 63, 52, 54, 36, and 51, who might be termed "one-type" readers, make a special effort to adapt their style of reading to the purpose in view. Similarly, on other occasions, special directions or suggestions were offered to an individual or to certain members of the group. As a further help, individual conferences were held outside of class with students Number 50, 28, 59, 63, and 58, in which their special difficulties and methods of correcting them were discussed.

Following the work described above, remedial instruction took the form of practice periods in paragraph comprehension. Each day, one of the nine specific objectives listed in the second major division of the outline was taken as the objective for that day. For example, the first day's work in paragraph interpretation was a practice period devoted to discovering the relevance of a series of paragraphs to a topic, and the second to finding the main idea in each of a series of paragraphs. In the tenth and final period in this unit of work, an informal test of paragraph interpretation was administered.

Twenty-five paragraphs were selected for each day's work, although for the most part only fifteen to twenty were covered during a class period. Paragraphs were selected which seemed best adapted to the type of interpretation

¹ Ferdinand Schevill, A History of Europe, Chapter XXIV, p. 511-27. New York: Harcourt, Brace and Company, 1925.

demand. For example, paragraphs for the first period's work were selected from Chapter II of the book by F. C. Dietz, The Industrial Revolution. There were no paragraph headings, the material was relatively simple, and paragraphs could be selected which contributed to the main idea or theme of the chapter in a variety of ways. For the second period, paragraphs were selected from Chapter IV of Beard's Whither Mankind. These paragraphs were also without marginal or side heads - an important consideration in view of the objective mentioned in the previous paragraph. Material for the third period was selected from Chapter XXIV of Schevill's History of Europe. In this instance the excellent marginal and side heads were essential. And so from day to day paragraphs were chosen to give a particular kind of practice.

During this phase of the work - and in fact throughout the remainder of the course - materials were not always selected which bore directly on the topic under consideration in the history class. A considerable part of the material used pertained to the work already covered, that is, to the unit of history work mentioned above. It seemed highly important to select paragraphs adapted to the purposes of the reading class whether they related to the discussions current in History 141 or to topics already covered.

Each practice period was opened by a brief discussion of the objective selected for the day and of the manner in which the paragraphs would need to be read in order to achieve that purpose. The class then turned to the page indicated by the instructor and located the paragraph which he indicated by some such statement as, "Find the second paragraph on page 87, beginning with the words 'As the Gothic hordes ...'" The class then set to work to read the paragraph silently in order to comprehend it just well enough to respond in the manner indicated. As soon as a student was ready to respond, he closed the book - holding the place with pencil or marker - and waited quietly for others to finish. In so far as possible, all were allowed to finish reading. However, differences in rate of work were so great that the slowest readers could hardly be allowed to finish in every instance. The

instructor tried to give the slow reader an opportunity and also to maintain the interest of the rapid reader. One or more students were asked to give their interpretations of each paragraph and to comment on the interpretations given by others. The instructor then indicated the location of the next paragraph to be used, and the foregoing process was repeated. Since suitable consecutive paragraphs could seldom be found, the instructor tried in each case to bridge the gap from one paragraph to the next with an appropriate transitional statement. Thus, although used to some extent in isolation, the paragraphs were not entirely dissociated or lacking in continuity of thought.

Reference books could hardly be used in the practice periods devoted to sentence interpretation, the third phase of the remedial work. Too much time would have been lost in locating the sentences. Therefore, lists of sentences were selected from the references and mimeographed for class use. No effort was made to select sentences relating to a single topic. Nor did each list provide practice with a single type of sentence interpretation. Each sentence was selected with one of the specific objectives in mind (and this was indicated in the margin of the copy used by the instructor) but each list contained sentences requiring many types of interpretation.

At the beginning of each class period in which sentence interpretation was the objective, the mimeographed list of sentences was distributed. Students were requested not to read any sentence until told to do so. The instructor then stated that the group would read the first sentence for whatever purpose it had been selected - perhaps, to get the meaning of a long, involved sentence well enough to restate the idea. The students sought to acquire the required meaning as soon as possible. They were told to look up as soon as they were ready to respond. While reading one sentence, the students covered the sentences following it with a plain white, five-by-eight card. After hearing responses concerning one sentence, the instructor told the class what to do with the next sentence, and so on. A sentence test was administered in the last period devoted to sentence study. The time not required for taking the test was spent in discussing the students' responses.

Word meanings were dealt with in practically all class periods. It will be recalled that "the development of a larger reading vocabulary" was recommended as a corrective measure for about three-fourths of the students in the remedial group. Consequently, the problem was attacked incidentally from day to day. A list of the troublesome words encountered in the reading class periods was constructed. When possible, the student was merely helped in deriving the meaning of a word from an analysis of it or from its use in the sentence where it was found. Students were also encouraged to keep a list of difficult words encountered in their outside reading together with their definitions and to review the list frequently until the word meanings were learned. The following words and phrases illustrate the general nature of the black list constructed in the reading class:

espionage	defunct
implacable	autonomy
atheism	indemnity
political agitation	status quo
deism	equilibrium
plebiscite	pervasiveness
kaleidoscopic changes	longevity
coalition	incompatible
subsidized	amicably
cordon	neo-mercantilism

In addition to such incidental work in word meanings, four class periods were devoted to the discussion of some of the more common prefixes, suffixes, and word stems. Mimeographed lists were prepared for class use, giving in each case the word element, its meaning, and a few words illustrating its use. Students were asked to select the more important items, to give the meaning of as many of the items as possible without referring to the meanings given, to give illustrations other than those listed, or to give the literal meaning of words by combining the parts.

Final measurements.- The final class period on May 29, 1931 was used to administer reading tests comparable to those given during the first session of the remedial class, April 6, 1931. The test battery in this instance was made up of Form 4 of the Test of Achievement in Silent Reading, Form B of the Test of Reading Vocabulary and Form 200 of the Test of Accurate Interpretation in Reading.

In addition, the eye-movements of students in the remedial class and of the ten students in the control group were again photographed. This final measurement was based on a passage designated as Paragraph VIII, a paragraph equated as accurately as possible with Paragraph II of the earlier laboratory work. In equating the paragraphs, thirty people read them as though they were taking an individual test. A stop watch was used to secure accurate records of the time required to read each one. The averages for Paragraphs II and VIII were approximately the same. To evaluate the paragraphs further, six students taking no other part in the investigation were photographed while reading them. The data presented in Table XLIX show that their performance was much the same on both. Whatever advantage

TABLE XLIX

EYE-MOVEMENT RECORDS OF SIX COLLEGE FRESHMEN ON TWO
PARAGRAPHS SUPPOSEDLY OF EQUAL DIFFICULTY

Student	Number of Fixations		Number of Regressive Movements		Average Duration of Fixations	
	II	VIII	II	VIII	II	VIII
a	10.4	11.1	1.7	1.4	8.3	7.5
b	7.7	6.6	.6	1.8	6.6	5.6
c	9.4	11.4	1.1	1.2	7.0	7.3
d	9.5	8.6	.9	2.2	5.9	5.5
e	5.6	6.5	.9	.9	5.8	6.0
f	5.1	5.5	.6	.9	6.4	6.3
Mean	7.9	8.3	1.0	1.4	6.7	6.4

there was seemed to be in favor of Paragraph II, that is, it seemed to be slightly easier than Paragraph VIII. The procedure used in re-photographing the eye-movements of students in the remedial class and of students in the control group was identical with that in the original study except that only the trial paragraph and Paragraph VIII were used. The same slides bearing the directions were used in both instances.

¹
Paragraph VIII was also printed in 11-point type with lines four inches in length.

Results of the Experiment

Summary of test scores. - The test scores of the students on the initial and final reading tests mentioned in the foregoing section of this report are summarized in Table L. With the exception of one student, Number 51, all made higher scores on one or more parts of the final test than on the initial one. The averages also indicate improvement. The average gain in rate was a substantial one, bringing the mean for this group nearly to the mean for Chicago freshmen. Only three students scored lower than 3.00 on the final test. Small gains in comprehension and vocabulary are also indicated by the group averages.

Summary of eye-movement data. - The general improvement indicated by the test scores just presented is supported by eye-movement records obtained at the close of the training period. Table LI presents the eye-movement records of the students on Paragraph II, obtained during the winter quarter, and those on Paragraph VIII,

TABLE L
INITIAL AND FINAL READING TEST SCORES FOR STUDENTS
IN THE REMEDIAL READING CLASS

Student Number	Rate		Comprehension		Vocabulary	
	Initial	Final	Initial	Final	Initial	Final
50	3.27	3.89	41	38	14	18
55	3.01	3.89	45	53	25	30
59	4.38	4.70	36	35	18	15
63	2.44	2.66	28	44	18	15
52	3.44	3.78	46	49	24	18
54	2.61	3.78	56	64	25	35
48	3.13	5.30	43	39	21	21
47	3.07	3.27	49	55	23	20
37	3.58	4.75	51	50	20	21
61	3.17	3.27	42	53	21	21
60	2.61	2.87	53	49	21	22
36	2.66	3.46	51	51	29	38
51	2.44	2.14	54	50	22	19
58	3.62	4.18	50	52	21	26
62	3.33	3.97	40	48	26	31
28	4.45	3.62	45	45	16	24
Mean	3.20	3.72	45.6	48.4	21.5	23.4

used as a final measurement of progress. The class averages were better in every respect for Paragraph VIII; that is, there were fewer fixations, fewer regressive movements, and the average duration of fixations was shorter. Since Paragraph VIII was slightly more difficult than Paragraph II when the two were tried out together, these averages indicate improvement in the mechanics of reading.

TABLE LI
THE EYE-MOVEMENT RECORDS OF SIXTEEN STUDENTS BEFORE
AND AFTER A PERIOD OF TRAINING

Student Number	Number of Fixations		Regressive Movements		Duration of Fixations	
	II	VIII	II	VIII	II	VIII
50	10.7	9.0	1.7	3.0	6.2	8.2
55	8.8	6.6	.9	.5	9.4	7.0
59	7.9	7.9	2.2	2.2	7.3	7.5
63	10.4	11.9	3.6	4.5	7.5	7.9
52	10.1	9.2	1.6	1.2	6.7	7.4
54	9.6	5.0	2.0	.7	7.8	6.1
48	8.3	5.9	1.1	.7	8.0	6.7
47	10.0	9.8	2.2	2.0	8.2	7.7
37	9.4	6.1	.8	.8	9.2	7.2
61	8.1	8.3	.9	.7	6.6	6.8
60	10.8	10.3	2.8	3.6	7.5	6.8
36	7.8	6.3	1.7	.7	8.5	6.9
51	13.1	11.8	2.0	2.3	9.4	9.0
58	9.7	7.4	2.6	1.9	6.5	5.5
62	6.2	4.3	.4	.1	7.5	6.8
28	6.4	7.8	.2	.1	6.8	7.1
Mean	9.2	8.0	1.7	1.6	7.7	7.2

Achievement of individual students. - Since the remedial work was by no means equally beneficial to all members of the class, its effectiveness should be considered in terms of the progress of individual students as well as in terms of group averages. Moreover, an individual's gain or loss should be considered in relation to the nature of his difficulties. For example, if a slow rate of reading was the outstanding symptom of difficulty, correct remedial instruction should have resulted primarily in improvement in reading rate. Table LII is presented to show

the gains and losses of individual students and relate them in a general way to the difficulties involved. It constitutes a summary of the progress, or lack of progress, for each student in the remedial class as indicated by the test scores and eye-movement records given in Tables L and LI. The students are again placed in four groups corresponding to those used in diagnosis.

Two of the students classified as "low in all phases of reading", Numbers 55 and 54, showed improvement in all phases of reading by every measure employed.

TABLE LII

GAINS AND LOSSES OF STUDENTS IN THE REMEDIAL READING CLASS
ACCORDING TO THEIR FINAL TEST SCORES AND EYE-MOVEMENT RECORDS

Student Number	Amount of Increase* or Decrease* in:					
	Rate Score	Compr. Score	Vocab. Score	No. Fix. Per Line	No. Repr. Movements	Duration of Fixations
Low in All Phases of Reading:						
50	.62	3*	4	1.7	1.3*	2.0*
55	.88	8	5	2.2	.4	2.4
59	.32	1*	3*	0.0	0.0	.2*
63	.22	16	3*	1.5*	.9*	.4*
52	.34	3	6*	.9	.4	.7*
54	1.17	8	10	4.6	1.3	1.7
Satisfactory Rate:						
48	2.17	4*	0	2.4	.4	1.3
47	.20	6	3*	.2	.2	.5
37	1.17	1*	1	3.3	0.0	2.0
61	.10	11	0	.2*	.2	.2*
Satisfactory Comprehension:						
60	.26	4*	1	.5	.8*	.7
36	.80	0	9	1.5	1.0	1.6
51	.30*	4*	3*	1.3	.3*	.4
58	.56	2	5	2.3	.7	1.0
Good Mechanical Habits:						
62	.64	8	5	1.9	.3	.7
28	.83*	0	8	1.4*	.1	.3*
Number of Students Making Gains						
	14	8	9	12	10	10

*The asterisk is used to indicate a decrease in the case of test scores and an increase in the case of eye-movement records. In both cases, the asterisk identifies records indicating poorer performance in connection with the final measurements than in connection with the earlier ones.

The progress of Number 54 was greater than that of any other student. At the time the remedial work began, she was one of the poorest readers in the group. When it closed, she was the most efficient reader in the remedial group and one of the better readers in the entire freshman class. Number 55 also registered large gains: a rate of reading nearly one-third more rapid than that employed on the pre-test, larger comprehension and vocabulary scores, and eye-movement records indicating greater fluency. Other students "low in all phases of reading" showed improvement in certain aspects of reading. The gain in rate and vocabulary by Number 50, that in rate and especially in comprehension by Number 63, and that in rate and comprehension by Number 52 should be noted. Their eye-movement records give additional evidence of the gains indicated by the test scores of students in this group.

The remedial instruction was relatively less successful with the four students having "satisfactory reading rate" than with the group mentioned above. Two of them made no higher comprehension scores on the final test than on the pre-test. None made a higher vocabulary score. On the other hand, Numbers 47 and 61 made higher comprehension scores and, in spite of the fact that rate was not stressed with these students, all made higher rate scores and three of them better eye-movement records in the case of the final measurements.

Three of the four students having "satisfactory comprehension" made higher rate and vocabulary scores and all had better eye-movement records in at least two particulars. The gains of Numbers 36 and 58 especially should be noted. Their improvement is registered at exactly the points where, according to the preliminary diagnoses, the needs were greatest.

One of the students with "good mechanical habits," Number 62, showed improvement in every phase of reading. The other, Number 28, showed little improvement save in vocabulary development.

Eight students, Numbers 55, 53, 48, 47, 37, 36, 58, and 62, had final eye-movement records distinctly superior to those obtained in the original measurement. The final row in Table LII shows that fourteen students - seven-eighths of

the group - increased their rates of reading, and that eight to twelve students - half to three-fourths of the group - showed improvement on every other measure included in the table.

Comparison of the experimental and control groups. - Further evidence of progress among students in the remedial group is found in a comparison of the achievement of ten of them with the achievement of the ten students in a control group. The manner in which the cases were paired has already been explained.¹ Similarly, the basis of comparison has been set forth.² In Table LIII the eye-movement records of the two groups are presented; for Paragraph II, the initial measurement, and Paragraph VIII, the final one.³

TABLE LIII

NUMBER OF FIXATIONS AND NUMBER OF REGRESSIVE MOVEMENTS PER LINE
FOR TEN PAIRS OF COLLEGE FRESHMEN IN AN INITIAL AND FINAL MEASUREMENT

Control Group					Ten Remedial Cases				
Student Number	Fixations		Regr. Movements		Student Number	Fixations		Regr. Movements	
	II	VIII	II	VIII		II	VIII	II	VIII
53	8.7	10.7	2.3	3.2	51	13.1	11.8	2.0	2.3
12	11.0	13.4	2.4	2.5	60	10.8	10.3	2.8	3.6
32	12.3	8.8	2.6	1.6	63	10.4	11.9	3.6	4.5
31	10.7	13.2	2.5	3.9	54	9.6	5.0	2.0	.7
38	10.4	13.1	1.7	3.6	58	9.7	7.4	2.6	1.9
3	8.9	7.4	2.0	1.5	55	8.8	6.6	.9	.5
30	8.4	10.2	1.3	2.8	37	9.4	6.1	.8	.8
40	7.3	10.0	1.0	1.7	36	7.8	6.3	1.7	.7
15	5.1	5.5	.6	.9	62	6.2	4.3	.4	.1
16	5.6	6.5	.9	.9	28	6.4	7.8	.2	.1
Mean	8.8	9.9	1.7	2.3	Mean	9.2	7.8	1.7	1.5
Mean for the Entire Remedial Class						9.2	8.0	1.7	1.6

¹See pp.122-3, especially Table XLIV, p. 123.

²See p. 150.

³Since only slight differences were found between the duration of fixations of good and poor readers in the laboratory work reported in Chapter III, this factor has been omitted from consideration in this comparison.

The averages given in Table LIII indicate that, in general, students in the remedial group made greater improvement in the mechanics of reading than those in the control group. The former averaged 1.4 fewer fixations and 0.2 fewer regressive movements per line on Paragraph VIII than on Paragraph II, whereas the latter averaged 1.1 more fixations and 0.6 more regressive movements per line for Paragraph VIII.

The differences noted above are exactly opposite in character to those which probably would have occurred had no remedial instruction been given. The control group had a slight advantage over the remedial group with respect to reading ability according to the original measurements. Moreover, since the control group ranked considerably higher with respect to native ability, it seems reasonable that they should have made greater progress in reading as a result of general reading requirements in college than the students in the remedial class. Since they did not, it seems highly probable that the improvement shown by the eye-movement records of students in the remedial group should be attributed largely to the remedial instruction which they received. Had it not been given, it seems altogether unlikely that their records would have surpassed those of slightly more capable students to the extent of 2.1 fewer fixations and 0.8 fewer regressive movements per line when reading Paragraph VIII.

Two additional facts strengthen the evidence presented above. First, since Paragraph VIII was slightly more difficult than Paragraph II according to the data obtained in the preliminary comparison of them, the eye-movement records of the control group on the two paragraphs bear the relation which one would expect for students whose reading habits were unchanged. Second, the ten students from the remedial class used in this comparison are quite representative of the remedial class as a whole. The averages for the entire class, given in Table II and also in Table LIII, are almost identical to those of the ten students paired with members of the control group. Moreover, the sample included some of those who profited least from the remedial work, for example, Students 51, 60, 63, and 28, as well as those making satisfactory progress.

Summary of results. - Since the accuracy of the student gains enumerated in the foregoing sections depends on the accuracy of the measures used, the possibility of overrating a single bit of evidence has been kept in mind. The analysis has been directed toward an evaluation of progress, not in terms of any one measure, but in accordance with all the evidence at hand. The data in Table LIV provide a general overview of the results of the remedial work - a summary of the data presented above. In the case of four students, all the evidence points toward definite improvement in all phases of reading. Eight students made moderate gains, ranging from slightly higher scores on two parts of the final tests to relatively large gains on five of the six final measures. The range of improvement among students in this group may be noted by contrasting the records of Students 60 and 36 in Table LII. Finally there were four students who apparently made little or no improvement in reading as a result of the remedial instruction.

TABLE LIV

PER CENT OF STUDENTS IN THE REMEDIAL CLASS WHO GAINED
IN VARIOUS PHASES OF READING

Student Number	Indication of Improvement	Per Cent of Cases
55, 54, 58, and 62	Better records on all final measurements.	25
50, 63, 52, 47, 37 61, 60, and 36	Better scores on two parts of the final reading test, and eye-movement records that surpassed the earlier ones in one or more particulars. ^a	50
59, 48, 51, and 28	No final test score, or only one, that was higher than that on the pre-test. ^b	25

^a In the case of Number 63 there was no improvement in eye-movement records.

^b Numbers 48 and 51 had better eye-movement records on the final measurement than on the earlier one; Numbers 59 and 28 did not.

The discussion now turns to the meaning of these results and to a consideration of some of the factors which help to explain the progress, and lack of progress, apparent from the foregoing presentation.

Evaluation of the Experiment

The value of the plan of remedial instruction developed in this study is attested by the improvement in reading habits cited in the preceding section of this chapter. On a series of final measurements, all class averages were superior to those derived from comparable initial measurements; large gains were indicated by the scores of certain students; in other cases, a cumulation of small gains seemed to indicate definite, though less extensive, improvement. Progress in various amounts was achieved in three-fourths of the cases involved. Moreover, since a control group of slightly more capable students showed no improvement in reading without instruction, the gains made by students in the remedial class were apparently due to the instruction given. Although the achievements of certain students in the group were far from satisfactory, the results obtained justify further experimentation with the remedial procedures followed in this study.

The achievements of certain students indicate that the procedure had its points of strength. The lack of progress or limited progress on the part of other students indicates that the plan should be improved. Since the advantages of the method have been set forth already in earlier sections of this chapter, further discussion of its merits is unnecessary at this point. On the other hand, certain weaknesses should be pointed out. The paragraphs which follow offer suggestions for the improvement of the plan, based on observations of the way it functioned in the present investigation.

A suitable time for remedial work. - Some of the disadvantages of having to postpone the remedial program till the spring quarter have already been pointed out.¹ In addition to the administrative problems involved, there are

¹See p. 122.

other considerations scarcely less important. If a student has been able to meet requirements during two previous quarters without remedial instruction, he naturally raises the question whether he could not continue to meet them without assistance. Furthermore, by the third quarter in residence the student has developed quite a number of interests. He cannot assume a new obligation of any considerable importance without giving up some activity. In other words, he has a full program and must sacrifice some active interest in order to participate in any extra-class activity such as a class in remedial reading. Even weather conditions are somewhat less favorable in the spring quarter, for it is not especially easy to remain indoors doing an "extra" piece of work when the first pleasant days of the season arrive. First of all, therefore, experience in this instance indicates that whatever remedial instruction is given should begin immediately upon the student's entrance into college.

Selection of cases. - Organization of the remedial class on a volunteer basis was relatively unsatisfactory in this investigation. It seems evident that participation in remedial work should be determined by a student's needs rather than by his inclinations. The fact has already been mentioned that the remedial group organized in connection with this study did not contain the sixteen poorest readers in the 1930-31 freshman class, partly because of conflicts in students' programs, partly because of the withdrawal of the poorest readers, and partly because of lack of interest. If remedial work is to serve its intended purpose, it should be given to the students whose reading difficulties constitute the most serious handicap. To do so in a class of the type used in this investigation, the work should be given during the first quarter's residence and be organized in connection with a course required of all poor readers, at least. If facilities are available, students of less than average ability (but not remedial cases) might conceivably be allowed to take part, as they were in this study. However, the program should be designed for, and carried out in the interests of, students who are handicapped most in reading.

Differences in the amount of improvement made by different students in the remedial class suggest an important question which this study does not answer: Is it possible to determine in advance the students who can, and those who cannot, profit by a particular kind of remedial instruction? If it were possible to identify the causes of reading disability which make the removal of deficiencies virtually out of the question - if there are such causes - then the recommendation made in the foregoing paragraph should be amended to read: remedial instruction should be provided for those who need it most and will profit most from it.¹

Extent of the training period. - In the present investigation the period of training lasted only eight weeks. Except in the case of the first unit of work, which dealt with the interpretation of entire selections, the time available did not seem adequate for the task at hand. Additional periods devoted to practice in word building seemed especially desirable. Better results undoubtedly could have been obtained had it been possible to extend the period of actual instruction to twelve weeks, even though the procedure used was the same in every other particular. For some students instruction for two quarters, or even a whole year, might be advisable.

Materials of instruction. - The desirable features of the reading material used in History 141 have been pointed out - the quantity, the variety, its immediate value to the student, and the use that students were expected to make of it. However, these advantages were offset, to some extent at least, by certain disadvantages.

First of all, there was the matter of the student's natural interest. Those who liked history and were taking it without anyone's suggestion found the materials interesting. Those who took the course on the advice of the remedial worker or the dean were interested in the materials as a factor in meeting

¹It is fully recognized that difficulties which cannot be corrected by one method may yield before another, precluding any policy of rigid determinism. Nevertheless, a knowledge of the difficulties that cannot be corrected by a particular type of instruction would eliminate unnecessary waste in remedial work.

course requirements; but not always because of the topic under discussion. This was not conducive to the best class work. Moreover, because of this fact, the remedial worker could make but limited use of the material as a lead to independent reading.

A second difficulty often encountered was that of finding materials appropriate both to the objective of the forthcoming practice period and to the topic under discussion in the history class. This was especially true when the work in reading was concerned with paragraph and sentence interpretation. The remedial worker was often forced to select materials related to any phase of the work already covered in history instead of those related to the work at hand. Since that was true, there was little reason why the materials for the later units should have been drawn from History 141. The appeal of immediate interest was gone. The periods were no longer devoted to reading selections which constituted part of the regular daily assignments. Incidentally, this situation was probably responsible to some extent for the fact that absences increased in number after the completion of the first unit of work.

Closely related to the foregoing problem was the matter of keeping the work of the reading class from being submerged. When the class was confronted with an important assignment in history and the students and their history instructor emphasized its importance to the remedial worker, it was rather difficult to set aside that material no matter how ill-adapted it was to the proposed work in reading.

Finally, there was the problem of the difficulty of the material. In the opinion of the remedial worker, one of the most serious defects in the remedial work was the use of materials too difficult for the students concerned. Every effort was made to select the most simple passages available which were appropriate in view of other considerations, especially during the first months' work. But although the books recommended for the course varied greatly in form and content, they contained all too few easy selections. One could have gone outside the recommended list, but there were two objections to doing so:

(1) students would have had to purchase the necessary books, and (2) since they were expected to read the books on the recommended list, the use of other material in the reading class would have been regarded by them as an additional requirement.

One criterion of the difficulty of material is the vocabulary used in it. Data are presented in Table LV showing the percentage of easy and difficult words in sample passages from six of the books used in the course.¹ At first glance it seems that the percentage of easy words is high. If 89 to 93 per cent of the words in a selection are found among the 5,000 most commonly used words, or at least are derived from those words, one might be inclined to discount vocabulary troubles. But in earlier phases of this study evidence was given that a considerable number of words in the upper 5,000 of Thorndike's list are not in the reading vocabulary of college freshmen. Moreover the final column of the table shows that about 5 per cent of the words do not belong even among the 10,000 most commonly used words. In view of the limited vocabularies of students in the remedial class, it seems probable that about 5 per cent of the words encountered were somewhat troublesome even to the better readers in the remedial group. In that case, vocabulary difficulties would arise on the average about once in every line and a half. Anyone who has recently attempted to read material where words with which he is unfamiliar occur as often as one in twenty will recognize the importance of this factor. Other causes of difficulty, such as, involved sentences, unusual expressions inadequate, explanation, poorly chosen illustrations, and the omission of transitional words and phrases cannot be determined objectively in so far as the materials of this course are concerned. Nevertheless, students' errors in connection with paragraph and sentence interpretation revealed the fact that

¹ The method used was not as thoroughgoing as those which have sometimes been employed, but the results show clearly enough the importance of vocabulary difficulties in connection with this project. For description of more refined methods see: Bertha Lively and S. L. Pressey, "A Method for Measuring the 'Vocabulary Burden' of Textbooks," Educational Administration and Supervision, IX (October, 1923), 389-98 and Alfred S. Lewerenz, "Measurement of the Difficulty of Reading Materials," Los Angeles Educational Research Bulletin, VIII (March, 1929), 11-16.

TABLE LV

PER CENT OF EASY AND DIFFICULT WORDS FOUND IN
SAMPLE PASSAGES FROM SIX REFERENCE BOOKS

Books	Length of Pas- sage	Begin- ning on Page	Per Cent in Thorndike's 1st 5,000 words	Per Cent Not in the 1st 5,000 Words	Per cent Not in the 1st 10,000 Words
Randall, <u>Making of the Modern Mind</u>	2000	(--) ^a	89.8	10.2	4.7
" "	300	320	89.3	10.7	5.3
" "	500	320	89.6	10.4	4.6
Schevill, <u>History of Europe</u>	500	356	89.0	11.0	6.2
Dietz, <u>Industrial Revolution</u>	500	54	92.2	7.8	3.4
Beard, <u>Whither Mankind</u>	500	204	92.6	7.4	4.0
Dickinson, <u>International Anarchy</u>	500	239	93.0	7.0	4.2
Clark, <u>Great Short Biographies of the World</u>	500	702	92.8	7.2	3.8

^a First of all, a word count was made in Randall's book using the first hundred words on pages 50, 100, 150, 200, 300, 350, 400, 450, 500, and 600. In this sample 89.2 per cent of the words were found in the first 5,000 words of Thorndike's Word List. Then turning to page 320, the middle of the book, another 1,000-word sample was checked against the list and combined with the foregoing sample to give the percentage 89.8, which is reported in Table LV. Separate records kept for 300-word and 500-word samples on page 320 gave approximately the same result; so, in checking the five other books 500-word samples were used, taking 500 consecutive words at the middle of each volume. Following Thorndike's own plan, plurals, adverbs formed by adding ly, participles and adjectives formed by the addition of a final n, comparatives and superlatives, and verb forms in s, d, ed, and ing were counted as belonging in the list if the word from which they are derived is in it.

difficulties due to style and organization were frequently encountered.

Attendance. - In the present experiment attendance was placed on a voluntary basis, just as was participation in every other phase of the investigation. The irregularity in attendance which resulted from this arrangement was a serious hindrance to successful remedial work. Absences ranged from two days in the case of student Number 36 to twenty days, or half of the class sessions, in the case of Number 52. The average for the class was 9.5 days' absence, or nearly one-fourth of the total time. Those who were least regular in attendance showed least interest in the remedial work when present and the smallest amount of improvement as a result of it. The data in Table LVI provide ample evidence of the

relation between regularity in attendance and the amount of improvement shown. In this table, the students are grouped in the same manner as in Table LIV, that is, according to their progress.

TABLE LVI
NUMBER OF DAYS' ABSENCE FOR STUDENTS IN THE REMEDIAL CLASS IN
RELATION TO THE PROGRESS SHOWN

Degree of Progress	Student Number	Number of Days Absent	
		Each Student	Average
Gains	55	7	
Indicated	54	5	
by All	58	9	
Records	62	4	
Average for the Group			6.3
Moderate	50	10	
Gains	63	7	
Indicated	52	20	
by	47	7	
Certain	37	13	
Records	61	3	
	60	12	
	36	2	
Average for the Group			9.3
No	59	16	
Improvement	48	15	
Shown	51	13	
	28	9	
Average for the Group			13.3
Mean for the Entire Class			9.5

Attendance records of the first group stand in striking contrast to those of the group who showed no improvement. Obviously, the lack of progress of Students 59, 48, 51, and 28 should not be attributed solely to defects in the method of instruction that was employed. Experience in this investigation seems to indicate that attendance in remedial classes should be compulsory rather than voluntary, enforced by whatever administrative sanctions may be necessary.

Provision for individual differences. - The plan of remedial instruction under discussion does not, within itself, provide for individual differences in interests, needs, and abilities. It has a certain degree of flexibility, but it tends to invite uniformity. In this investigation the attempt was made to provide for individual differences in the following ways: (1) by placing major emphasis on the development of abilities needing improvement in the case of the greatest number of students through the choice of objectives and the distribution of class time, (2) by directing the attention of students to their own needs and to desirable corrective measures, (3) by providing supplementary material for classroom use by the more rapid readers in the group, (4) by suggesting materials for independent reading which seemed appropriate to the abilities and interests of the student, and (5) by conferring with individual students on reading problems. But in spite of these efforts, it is evident that further differentiation was needed. Each case in which the student failed to become interested in the work, and to profit by it, reveals the need for further individual attention.

Summary. - Since three-fourths of the students in the remedial class showed some degree of improvement in silent reading achievement under the plan of instruction used in this study, the method has important possibilities. The remarkable progress of certain students, including some whose reading habits were quite immature, is further evidence of the potential value of this type of instruction. Although all the reading difficulties of students in the remedial class were by no means eliminated, the results seem to warrant further experimentation with the method. In future experiments, however, the following modifications seem desirable: (1) conduct the class during the students' first quarter of residence; (2) give instruction only to those whose reading difficulties constitute a serious handicap; (3) let the period of actual instruction extend through the entire quarter; (4) organize the work in conjunction with a course in which there is an abundance of easy reading material; (5) make attendance compulsory; and (6) adapt the work as fully as possible to the interests, needs, and abilities of the individual student.

CHAPTER V

INTERPRETATION

This chapter deals with the educational significance of the investigation described in the foregoing pages, a study concerned with the measurement and improvement of reading among college freshmen. Since the findings relating to the reading abilities of college students are of basic interest in the study, these are considered first. Following the interpretation of reading achievement as revealed by this inquiry, the discussion turns, first, to an evaluation of the work in test construction, and then, to the significance of the project in remedial instruction. The chapter closes with the suggestion of important problems upon which further research is needed.

Evaluation of the Findings Pertaining to Reading Ability

One of the significant contributions of this investigation is a body of facts relating to the reading achievements of college freshmen. Reading habits are described in accurate, objective, and meaningful terms: first, by means of test scores; then by eye-movement records. The facts presented make possible a clearer understanding of the manner in which students read at the time they enter college. Data presented in Chapter III provide tentative standards of normal achievement; they show the range in ability that exists and how ability is distributed within a group; they reveal the extent to which serious reading difficulties are present; they suggest minimum acceptable standards; and, they contribute to an understanding of the way in which specific abilities are interrelated. Each of these factors is important: in administration, in guidance, and especially in college teaching.

Standards of achievement.- By contributing a meaningful standard of reading rate, that is, a standard expressed in terms of words per second, this study provides the college teacher with a criterion for his reading assignments. Knowledge of the normal reading rate of the average student is

necessary to a sense of proportion in teaching. Without it, reading assignments are likely to be unreasonably long or so brief that they offer little challenge. Most college students hear suggestions such as these: "In this course you will be expected to read at least a page a minute; if you cannot, it will be hard for you to cover the required readings"; or, "Have a friend or classmate time you as you read; unless you turn a page every minute, you are reading too slowly." Note that the second illustration requires a reading rate twice as rapid as the first. Neither agrees with the data of this study. Only 25 of the 664 students in the 1931-32 freshman class read at the rate of a page a minute; that is, at the rate of six words per second (6.00) or 360 words per minute, a conservative estimate of the number of words per page. In so far as the data of this study are dependable, the reading rate of the average freshman student, four words per second (4.00), is considerably slower than the rate frequently advocated. The discrepancy is not surprising. Teachers interested in subject matter more than in students have accepted uncritically various so-called standards. Those who were sincerely interested in the study habits of their students but handicapped by a lack of accurate measuring instruments have often found no better standard for their students than the rate at which they themselves read the materials in the field in which they have specialized. The rate standard suggested by the data of this study, verified as it is by repeated measurements with reading tests and by eye-movement records, should enable college teachers to establish reading requirements in harmony with the ability of their students.

Test scores of the average college freshman on the vocabulary and comprehension parts of the test battery developed in this study have somewhat less practical significance as independent standards than the rate score. Like the norm on any point-score test, however, these standards are meaningful when

related to the tests on which they are based.

Typical reading habits, as portrayed by the eye-movement records obtained in this investigation, deserve the attention of college teachers. College instruction is based frequently, perhaps in a majority of cases, on the assumption that most college students read in a fluent manner. The data of this study do not warrant such an assumption. The average Chicago freshman, in the group whose eye-movements were photographed, could hardly be termed a fluent reader when confronted with simple materials appropriate for mature students. Instead of reading with four or five short fixations per line and an occasional regressive movement, as is often assumed, this average student made 8.7 fixations per line, 1.7 regressive movements per line, and his fixation pauses averaged 7.5 twenty-fifths (or three-tenths) of a second. This finding has an important bearing on college education; but the college itself must interpret it. Perhaps it means that the primary interest of the college in the reading habits of its students is not remedial instruction, but rather the development on the part of the great majority of its students (the so-called average group) of reading habits similar to those of students who now rank as superior readers. If so, the college must provide a program of reading instruction for most of its entering students. A more logical inference seems to be that the college needs to recognize the present abilities of its entering students and to adapt its work to them, raising its standards only as improvement in public-school work results in student bodies with better reading habits. It is neither good pedagogy nor sound counsel to set up misleading standards for students to follow.

Knowledge of the typical reading ability of college students is useful outside the classroom. As the findings of this study are extended and refined, they should constitute an important guide: to administrators, in admitting and classifying students; to personnel directors, in advising students; and, to the writers of textbooks and those who select them, in providing materials appropriate

for college freshmen. Even the tentative standards presented here should be helpful.

Distribution of ability.— The data secured in this study provide objective evidence of the range and distribution of ability among students in various phases of reading. The teacher of college freshmen should be able to make more adequate provision for individual differences when he knows: (1) that, in the same class, one student may read at the rate of seven words per second with as high a degree of comprehension as another who proceeds at the rate of two words per second; (2) that one student attaches no meaning or an incorrect meaning to about four-fifths of the simple non-technical words such as those used in the Test of Reading Vocabulary whereas another student knows practically all such words; (3) that differences in ability to comprehend seem to be as marked as differences in reading rate and reading vocabulary; and (4) that one student may ordinarily make only four or five fixations of the eye per line, with one regressive movement in two lines, while another makes fifteen fixations per line, four or five of which are regressive movements. The difference in eye-movements just noted is greater than that which Buswell found between the eye-movements of the average high-school pupil and the average second-grade pupil in reading easy material.¹

A measure of dispersion is almost, if not quite, as significant as one of range. The standard deviation is particularly meaningful in the case of the rate score, indicating as it does that approximately two-thirds of the entering freshmen have reading rates ranging between three and five words per second.

Knowledge of both the range and distribution of reading ability is a prerequisite to suitable provision for individual differences. Methods of college teaching would probably be modified considerably were every teacher fully

¹ Fundamental Reading Habits, p. 38, 41, and 43. The average second-grade pupil made ten or eleven fixations per line; the average high-school student, about six.

aware of the differences in reading ability within each of his classes. Since, in an entrance class carefully selected by the use of various other criteria, pronounced differences in reading ability were found, their presence among college freshmen generally should be anticipated and the importance of the differences clearly recognized. Were this study to result only in wider recognition of the fact that differences in reading ability are extensive enough to require differentiation in teaching methods, it would make an important contribution to higher education.

Extent to which serious difficulties are present.- The findings of this study confirm those of earlier investigations with reference to the fact that poor reading habits are a definite handicap to many college students. Since approximately 10 per cent of the students in both the 1930-31 and 1931-32 Chicago freshman classes were found to be seriously handicapped in reading, colleges should see more clearly the necessity for counteracting this defect in the preparation of their students. These findings should renew the challenge for every college to investigate the reading achievements of its own students and to provide an effective remedial program in accordance with the needs discovered.

Minimum standards.- Knowledge of the lowest standards which seem to be reasonably acceptable at the college level is quite important to the educational counselor, the remedial worker, the regular instructor who undertakes to aid his students in building up better reading habits, and the student seeking self-improvement. The tentative standards provided by the findings of this study should therefore be useful to teachers and students alike in directing attention to unsatisfactory reading habits. The test scores which, in this study, appear to be danger points below which the student should not fall (namely, a rate score of 3.00, vocabulary score of 20, comprehension score "Ct" of 40, and composite score of 90) can be established more definitely by extended analyses of the relation between the test scores and academic success.

Relation of abilities to one another.- The data dealing with the relation of

the various phases of reading ability to one another and those pertaining to the relation of reading ability to intelligence and to achievement suggest the extent of those relationships, but warrant no final conclusions. In some instances, they supplement the findings of other studies, for example, the low positive correlation between rate and comprehension (.10 to .27) and between reading rate and intelligence (.31); also the comparatively high correlation between intelligence and the factors: general reading ability (.70), comprehension (.57), and vocabulary (.72). Other apparent relationships, such as that between word knowledge and reading rate, word knowledge and comprehension, or reading ability and academic success, suggest lines of inquiry which should be carried out. From an accumulation of findings such as these with respect to the interrelationship of abilities, conclusions can ultimately be drawn for the guidance of case workers, teachers, and administrators. The policy of the college with respect to the whole matter of reading instruction and the procedure of regular and remedial teachers should eventually rest on a body of known facts about the specific abilities and their interrelationships.

Significance of the Work in Test Construction

General significance of the reading tests.- Since differences in reading ability among college students are highly important, each improvement in the technique of measuring that ability has potential value in the improvement of college organization, administration, and instruction. One of the contributions of this investigation is the development of an instrument of measurement, a battery of silent reading tests appropriate for use with college students. Although developed primarily for use in this investigation and not yet standardized, this battery of tests provides a dependable measuring instrument.¹ Without disregarding the limitations set forth in Chapter II, the following merits of the test battery may be

1

The reliability coefficients given in Chapter II, p. 60, are as follows: for the composite score, $.892 \pm .005$; rate score, $.747 \pm .011$; comprehension score "Ct", $.703 \pm .013$; and vocabulary score, $.908 \pm .005$.

pointed out. It yields a rate score that can be translated readily into words per minute, pages per hour, or other desired units. This score, based on the continuous reading of non-technical material written in language appropriate for the mature reader, is derived from a reading period long enough to assure maximum reliability. The vocabulary test is brief, highly reliable, and practically untimed, thereby eliminating the influence of reading rate from the vocabulary score. The comprehension test shows not only the general level of ability which the student possesses, but also the particular phase of comprehension in which the student is weak. Each item in the entire test battery has been validated by statistical methods. The validity of the test is further indicated by the close agreement between the test scores and eye-movement records of fifty students. There are four comparable forms. Finally, the test is not difficult to administer, score, or interpret. In many of these respects it represents an improvement over any test battery available. Experience with the test in this investigation warrants the assertion that with it a meaningful measure of reading ability can be obtained, one that can be used with considerable assurance in determining the general level of reading achievement in a group, in identifying poor readers, and in diagnosing reading difficulties.

Educational levels for which the test is appropriate.- Although designed especially for use with junior-college students, data obtained in the course of this investigation indicate that the test may also be used effectively with high-school seniors, senior-college students, and graduate students. The range of difference between mature groups at the various academic levels is apparently narrower than that of individual differences within a given group. Standards would undoubtedly vary to some extent, but the test should constitute a significant measure of the reading ability of mature readers generally. It may be used for three types of measurement: (1) general survey, (2) diagnosis, and (3) measurement of progress as a result of instruction.

The general survey.- A high-school counselor needs the kind of information about the reading ability of his high-school seniors which can be obtained with a

general survey test. He can give sound advice to students who are planning to pursue academic work beyond the high school only when he knows what reading abilities and disabilities they have. A general survey of reading ability is equally valuable; or perhaps more so, at the time the student enters college. It gives the administration insight into the most important study habits of the entering group and provides a knowledge of the differences in reading ability essential to the most effective classification of students. Guidance departments should find the results helpful with respect to both educational and vocational guidance. Instructors should know something of the reading abilities of their students in order that they may plan their work to better advantage and make adequate provision for individual differences. Finally, a general survey of reading ability helps to identify the poor reader, the student who is actually handicapped because of unusual disabilities. Through the survey, the need for a special program of remedial work - if such a need exists - is revealed.

Diagnosis.- Scores on the various parts of the test indicate the particular phase of reading in which the student is weak. Diagnosis must go beyond test scores; but the latter are none the less significant because of that fact. The test scores provided by this battery of silent reading tests guide the teacher to specific reading difficulties and often suggest their causes, especially when the students' errors are carefully analyzed. In diagnosis, important facts can often be discovered by giving the test as an individual exercise and observing the student as he works.

Measurement of improvement.- When a program of instruction is carried out looking to the improvement of reading habits, the measurement of progress is important. The four forms of this reading test make it possible to determine progress from time to time by repeated measurements.

Methods used in constructing the test.- The methods used in this investigation in developing plans for the tests, in selecting appropriate materials, in evaluating individual test items, and in comparing test results with the accurate records

of the laboratory should exert a small but important influence on the practices of test construction; not because these methods were original, but because ordinary procedures were used in new combinations and applied toward the solution of new problems. Likewise, some of the findings and experiences reported in this study with reference to test techniques and the characteristics that are, and are not, suitable for college reading tests should exert a wholesome influence on the reading tests developed in the future.

Evaluation of the Work in Remedial Reading

Need for remedial work.- The data of this study warrant the conclusion that serious reading difficulties are encountered by a sufficient number of college students to require definite remedial instruction more adequate than that which can be given in the regular college classroom. The need is demonstrated by both the number of students with poor reading habits and the large proportion of the poor readers doing unsatisfactory work in college. To the extent that the data of this study are typical, at least one-tenth of the students in an entering freshman class are seriously handicapped by poor reading habits. That is, for each hundred students in an entrance class approximately ten are unable to read well enough to do college work successfully. If they have other handicaps as well, such as undesirable traits of personality, physical handicaps, unfortunate attitudes and habits, or a limited background of experience, the reading handicap is especially serious. The findings of this and many previous studies point unmistakably to needless drudgery, to disappointment and discouragement, to the failure and subsequent elimination of students from college that follow in the wake of reading disabilities. The responsibility of the college in providing remedial instruction will be assumed only when college faculties realize fully the importance of reading problems. In pointing out the facts with respect to the extent of reading deficiencies among college students this study should stimulate interest in this vital educational problem. The scope of the study, the variety of measuring instruments employed, and the fact that Chicago freshmen are a selected group give added significance to the findings

presented.

Although knowledge of the extent of reading difficulties is especially important from the point of view of college administration, it is also an important consideration in college teaching. The college teacher cannot provide adequate remedial instruction for every student who needs it, but he can be on the alert for handicapped readers and can give them a considerable amount of guidance in reading. He is more likely to do so if he knows that in every unselected group of thirty students at least three are likely to be in need of individual guidance.

Possibilities of remedial instruction.- The extent to which the reading difficulties of mature readers can be corrected is a question only partially answered by the data of this study, for the faulty habits which persisted in spite of the training given might have yielded before other methods. However, large gains by certain members of the remedial class show that college students who are seriously handicapped in reading can often develop satisfactory ability, and varying degrees of progress on the part of three-fourths of the students provide evidence that remedial instruction has considerable value. More than a quarter of a million college freshmen enroll in American colleges each year. According to the most conservative estimate 25,000 of them are handicapped because of poor reading habits. If the colleges of the nation could help half of these students to develop better reading habits, that work would be exceedingly valuable, not only in preventing elimination from college, but also in increasing the personal satisfaction of students through increased efficiency, enriched experience, and better educational integration.

Value of the method used in this study.- The technique of remedial instruction developed in this study differs materially from any group method used previously, especially in the matter of organization and sequence of class activities. Although not equally beneficial to all students in the remedial class, the method was moderately helpful to a majority of them. The objectives chosen and devices employed seemed fundamentally correct. Progress seemed to be limited by certain

extrinsic factors more than by the method of instruction employed. Moreover, by revealing many of the conditions which limited the effectiveness of the work, this experiment in group instruction points the way to improvements in the plan developed which should increase its value. Further research will be needed to determine whether the suggested revisions make the plan entirely satisfactory and how it compares in effectiveness with other group methods. The findings of this study, however, are evidence of its potential value.

The objectives used in this experiment and many of the procedures employed should be suggestive not only to those interested in the continuation and improvement of this same general technique, but also to remedial workers who are using other plans of group or individual instruction. Since reading disabilities are so prevalent among college entrants, the development and evaluation of each new method of remedial reading is a step toward the solution of one of the most important current problems in college education.

Problems for Further Study

This investigation has touched upon a number of problems which it does not solve and raised a number of questions which it does not answer. Some of them are listed below together with other problems not attacked in this study but suggested by some phase of the work. The present investigation could make no contribution of greater importance than to stimulate a research interest in problems such as these.

Problems pertaining to the nature of the reading process.- None of the studies already made, nor all of them combined, give as thorough insight into the reading process as one needs in order to achieve maximum results in reading instruction. An analytical study of reading is greatly needed. Among the specific elements which should be singled out for intensive study, the following seem especially important. A thorough-going analysis of comprehension - the identification and evaluation of the basic elements which enter into it - would be particularly helpful in the teaching of reading. The findings of the present study and those of

Previous investigations with respect to the interrelationship among the various phases of reading ability should be extended and refined. Various factors which might conceivably affect reading habits should be singled out and their influence be measured accurately; for example, reaction time as a factor in the control of the muscles of the eye, or the influence of general physical condition on one's reading habits. In an intensive, analytical study of the type suggested here, the most refined techniques of measurement available should be employed. Such a study should result in a much better analysis of reading than is now available and provide a basis for more effective instructional procedure. Through it valuable standards could probably be determined. For example, optimum reading rate could be discovered for students at various levels with respect to factors such as intelligence, reaction time, and association time. The study would undoubtedly show the factors which promote and those which hinder the development of better reading habits, thus making it possible for the college to provide remedial instruction, first of all, for the students most likely to profit by it.

Closely related to the problem just outlined, but probably too extensive in scope to be included as a part of such an inquiry, is the thorough analysis of reading rate. Previous investigations have shown that reading rate varies greatly in the same individual, and, in a general way, the causes of such variations are known. But the relative influence of the many causal factors has not been accurately measured. Materials of many types, reading situations of various kinds, and a variety of reading purposes should be introduced under controlled conditions in order to determine the influence of each factor. An investigation of this type should result in a series of standards or norms and a series of minimum standards or danger points, similar in purpose to those given in the present inquiry but covering many types of reading situations.

An investigation of reading habits, similar to the one reported in Chapter III. of this report but using students in an unselected entrance class, would constitute a valuable supplement to this study. By administering reading tests to unselected students and using the "average student" instead of the superior reader and the

poor reader as subjects for the laboratory work, the findings of this study could be modified or extended. The investigation might well be broadened to include college students in the upper classes and graduate students as well as college freshmen.

Additional work in test construction.- Improved tests of reading comprehension are needed urgently. Further experimentation with tests modeled along the general lines of the Test of Accurate Interpretation in Reading might well be done, using materials of different types, validating individual items with a larger number of cases, and evaluating more carefully the various parts of the test. Further evaluation of the techniques of test construction reported in this study would also be distinctly helpful to test makers.

An investigation relating to remedial instruction.- The method of group instruction used in this investigation should be improved and tested further. This and various other group methods should then be evaluated through comparative use in controlled experiments to determine, not only the relative effectiveness of the methods themselves, but also the immediate and permanent effects of remedial reading on the student's reading ability, general study habits, interests, and achievements.

The problems suggested in the foregoing paragraphs imply a need for thorough study of specific problems. Based on the findings of investigations already reported, intensive studies of particular phases of reading should now be undertaken. Many of the techniques needed for such inquiries have already been developed. The need for applying them is urgent. Until colleges know more about the reading ability and reading habits of their students, little progress can be made in developing effective remedial reading programs. Further achievement in this important phase of college education depends in no small measure on the adequacy of future research on the reading problems of college students.

SELECTED REFERENCES

- Alderman, G. H. "Improving Comprehension Ability in Silent Reading," Journal of Educational Research, XIII (January, 1926), 11-21.
- Allbright, B. F., and Horning, F. "Typical Reading Disabilities of College Entrants," California Quarterly of Secondary Education, III (January, 1928), 166-69.
- Anderson, E. M. Individual Differences in the Reading Ability of College Students. University of Missouri Bulletin, XXIX (1928), Number 39. Education Series Number 25. Columbia, Mo.: University of Missouri, 1928. Pp. 77.
- Averill, L. A., and Mueller, A. D. "The Effect of Practice on the Silent Reading of Adults," Journal of Educational Research, XVII (February, 1928), 125-29.
- Barton, William A. Outlining as Study Procedure. Contributions to Education, Number 411. New York: Bureau of Publications, Teachers College, Columbia University, 1930.
- Behrens, Herman D. The Value of Remedial Reading for College Freshmen. Master's Thesis, Ohio State University, 1930. Pp. 196.
- Berry, B. T. A Study of Reading Comprehension of College Freshman Level. Doctor's Dissertation, University of Southern California, 1931. Pp. 245. See also: "Improving Freshman Reading Ability." English Journal, XX (December, 1931), 824-29.
- Book, W. F. How to Succeed in College. Baltimore: Warwick & York, Inc., 1928.
- Book, W. F. "How Well College Students Can Read," School and Society, XXVI (August 20, 1927), 242-48.
- Broom, Mybert Eustace. An Analysis of Certain Factors Affecting Reading Achievement at the College Level. Doctor's Dissertation, University of Southern California, 1931. Pp. 276.
- Buswell, G. T. An Experimental Study of the Eye-Voice Span in Reading. Supplementary Educational Monographs, No. 17. Chicago: University of Chicago, 1920.
- Buswell, G. T. Fundamental Reading Habits: A Study of Their Development. Supplementary Educational Monographs, No. 21. Chicago: University of Chicago, 1922.
- Carroll, R. P. An Experimental Study of Comprehension in Reading. Contributions to Education, No. 245. New York: Bureau of Publications, Teachers College, Columbia University, 1926.
- Carroll, R. P., and Jacobs, Charles G. "Drill in Silent Reading for College Freshmen," School and Society, XXX (November 9, 1929), 656-58.
- Crawford, C. C. "Some Results of Teaching College Students How to Study," School and Society, XXIII (April 10, 1926), 471-72.
- Crawford, C. C. The Technique of Study. Boston: Houghton Mifflin Company, 1928.
- Crofoot, B. L. "Remedial Reading," School and Society, XXV (February 12, 1927), 205-06.
- Dickinson, C. E. "A Study of the Relation of Reading Ability to Scholastic Achievement," School Review, XXXIII (October, 1925), 616-26.

- Dolch, E. W. Reading and Word Meaning. Boston: Ginn and Company, 1927. Pp 129.
Also: The Psychology and Teaching of Reading. Boston: Ginn and Company, 1931.
- Donovan, H. L. "Clinical Studies in Reading," Peabody Journal of Education, V (March, 1928), 269-77.
- Dransfield, J. E. "A Technique for Teaching Silent Reading," Teachers College Record, XXVI (May, 1925), 740-52.
- Eurich, A. C. An Experimental Study of the Reading Abilities of College Students. Doctor's Dissertation, Department of Education, University of Minnesota, 1929.
- Eurich, A. C. "A Method for Measuring Retention in Reading," Journal of Educational Research, XXIV (October, 1931), 202-08.
- Eurich, A. C. "The Photographic Eye-Movement Records of Successful and Unsuccessful College Students," Journal of Applied Psychology, XVII (October, 1933), 604-13.
- Eurich, A. C. "Relation of Speed of Reading to Comprehension," School and Society, XXXII (September 20, 1930), 404-06.
- Fitzsimmons, Ethel M. Study of the Ability of High School Seniors and College Freshmen to Read Directions. Master's Thesis, University of Nebraska, 1930.
Pp. 53.
- Gates, A. I. The Improvement of Reading. New York: The Macmillan Company, 1927.
- Gates, A. I. "What Do We Know about Optimum Lengths of Lines in Reading?" Journal of Educational Research, XXIII (January, 1931), 1-7.
- Germane, C. E., and Germane, E. G. Silent Reading. Chicago: Row, Peterson & Co., 1922.
- Gilliland, A. R. "Effect of Rate of Silent Reading on Ability to Recall," Journal of Educational Psychology, XI (November, 1920), 474-79.
- Good, C. V. "The Effect of Extensive and Intensive Reading on the Reproduction of Ideas or Thought Units," Journal of Educational Psychology, XVIII (October, 1927), 477-85.
- Good, C. V. "The Effect of Mental Set or Attitude on the Reading Performance of High-School Pupils," Journal of Educational Research, XIV (October, 1926), 178-86.
- Good, C. V. "The Effect of a Single Reading Versus Two Readings of a Given Body of Material," Journal of Educational Method, V (April, 1926), 325-29.
- Gray, W. S. "The Relation between Reading and Study," Addresses and Proceedings of the National Education Association, LVII (1919), 580-86.
- Gray, W. S. Remedial Cases in Reading: Their Diagnosis and Treatment. Supplementary Educational Monographs, No. 22. Chicago: University of Chicago, 1922.
- Gray, W. S. Studies of Elementary School Reading through Standardized Tests. Supplementary Educational Monographs, Vol. 1, No. 1. Chicago: University of Chicago, 1917.
- Gray, W. S. Summary of Investigations Relating to Reading. Supplementary Educational Monographs, No. 28. Chicago: University of Chicago, 1925.

- Greene, E. B. "Effectiveness of Various Rates of Silent Reading of College Students," Journal of Applied Psychology, XV (April, 1931), 214-27.
- Guiler, W. S. "Remedial Instruction," Educational Research Bulletin (Ohio State University), VIII (May, 1929), 265-66.
- Henmon, V. A. C. "An Experimental Study of the Value of Word Study," Journal of Educational Psychology, XII (February, 1921), 98-102.
- Hilliard, G. H. Probable Types of Difficulty Affecting Comprehension in Reading. University of Iowa Studies in Education, II (1918), No. 6. Iowa City: University of Iowa, 1918.
- Jacobs, Charles C. An Experiment in Silent Reading for College Freshmen. Master's Thesis, Syracuse University, 1929.
- Jones, E. S. Testing and Training the Inferior Freshman. University of Buffalo Studies, V (1927), No. 3. Buffalo: University of Buffalo, 1927.
- Judd, Charles H. Reading: Its Nature and Development. Supplementary Educational Monographs, II (1918), No. 4. (Whole number 10). Chicago: University of Chicago, 1918.
- Judd, Charles, and Buswell, Guy T. Silent Reading: A Study of the Various Types. Supplementary Educational Monographs, No. 23. Chicago: University of Chicago, 1922.
- Kathleen, Sister M. "A Comparative Study of Certain Silent Reading Tests," Catholic Educational Review, XXII (1924), 589-95; XXIII (1925), 33-42.
- King, Irving. "Comparison of the Efficiency of Slow and Rapid Readers," School and Society, VI (August 18, 1917), 203-04.
- King, Irving. "A Comparison of Slow and Rapid Reading," School and Society, IV, (November 25, 1916), 830-34.
- Lemon, A. C. An Experimental Study of the Guidance and Placement of Freshmen in the Lowest Decile at the State University of Iowa. Doctor's Dissertation, Department of Education, University of Iowa, 1926. See also: University of Iowa Studies in Education, Vol. III, No. 8. Iowa City: University of Iowa, 1927, Pp. 136.
- McGrath, Earl J. "The Three-Year Plan," Journal of Higher Education, II (June, 1931), 283-88.
- McNee, Marcia A. Factors Related to Reading Achievement of College Freshmen. Master's Thesis, Department of Education, University of Chicago, 1931. Pp 121.
- Merton, E. L. The Discovery and Correction of Reading Difficulties. Second Yearbook, Department of Elementary School Principals, National Education Association, 1923. Pp. 346-63.
- Miles, D. H. "Can the High-School Pupil Improve His Reading Ability?" Journal of Educational Research, XIV (September, 1926), 88-98.
- Miles, W. R., and Bell, H. M. "Eye-Movement Records in the Investigation of Study Habits," Journal of Experimental Psychology, XII (October, 1929), 450-58.
- Monroe, Marion. Methods for Diagnosis and Treatment of Cases of Reading Disability. Genetic Psychology Monograph, IV (November, 1928), Nos. 4-5, Pp. 343-456. Worcester, Mass.: Clark University Press, 1928.

- Morrison, A. F. The Improvement of Instruction in Reading through Diagnostic and Remedial Measures. Master's Thesis, Department of Education, University of Chicago, 1924.
- Morrison, K. A. The Relation between Reading Comprehension and College Success for First Semester Freshmen. Master's Thesis, Department of Education, University of Southern California, 1928.
- National Society for the Study of Education. Twenty-fourth Yearbook, Part I. Bloomington, Ill.: Public School Publishing Company, 1925.
- National Society of College Teachers of Education. "Quantitative Measurement in Institutions of Higher Learning." Eighteenth Yearbook. Chicago: University of Chicago Press, 1930.
- Nelson, M. J. "A Study in the Value of Entrance Requirements at Iowa State Teachers College," School and Society, XXXVII (February 25, 1933), 262-64.
- Odell, C. W. Predicting the Scholastic Success of College Freshmen. University of Illinois Bulletin, XXV (1927), No. 2. Urbana: University of Illinois, 1927.
- Parr, F. W. A Remedial Program for the Inefficient Silent Reader in College. Doctor's Dissertation, Department of Education, University of Iowa, 1929. See also: "The Extent of Remedial Reading Work in State Universities of the United States." School and Society, XXXI (April 19, 1930), 547-48.
- Parr, F. W., and Nemzek, C. L. "Inefficient Silent Reader in College," Peabody Journal of Education, VII (March, 1930), 299-303.
- Pitkin, Walter B. The Art of Rapid Reading. New York: McGraw-Hill Book Company, 1929.
- Pressey, Mrs. L. C. "College Remedial Reading Classes," English Journal, XIX (September, 1930), 566-69.
- Pressey, Mrs. L. C. "College Students and Reading," Journal of Higher Education, II (January, 1931), 30-34.
- Pressey, Mrs. L. C. A Manual of Reading Exercises for Freshmen. Columbus, Ohio: Ohio State University Press, 1928.
- Pressey, Mrs. L. C., and Pressey, S. L. "Training College Students to Read," Journal of Educational Research, XXI (March, 1930), 203-11.
- Quantz, J. O. "Psychology of Reading," Psychological Review. Series of Monograph Supplements, Vol. II, No. 1 (Whole number 5), December, 1897.
- Remmers, H. H. A Diagnostic and Remedial Study of Potentially and Actually Failing Students at Purdue University. Doctor's Dissertation, Department of Education, Purdue University, 1927.
- Remmers, H. H., and Stalnaker, J. N. "An Experiment in Remedial Reading Exercises at the College Level," School and Society, XXVIII. (December 22, 1931) 797-800.
- Robinson, F. P. "An Aid for Improving Reading Rate," Journal of Educational Research, XXVII (February, 1934), 453-55.
- Sangren, Paul V. Measurement of Achievement in Silent Reading. Doctor's Dissertation, Department of Education, University of Michigan, 1927.

- Schlichting, M. H. Methods of Improving Reading Achievement Among College and University Students. Master's Thesis, Department of Education, University of Chicago, 1931. Pp. 74.
- Schultz, Esther, and Miller, J. C. An Experiment Designed to Improve the Reading Ability of First Year College Students. Unpublished study. Columbia, Mo.: Christian College, 1928.
- Skinner, H. C. "Relationship between Reading Ability and Class Marks in College Subjects," Journal of Educational Research, XVII (February, 1928), 137-38.
- Stockrahm, R. L. A Tentative Analysis of Learning to Read with a Study of the Reading Ability of College Freshmen. Doctor's Dissertation, Department of Education, Indiana University, 1927.
- Stone, C. W. "Improving the Reading Ability of College Students," Journal of Educational Method, II (September, 1922), 8-23.
- Stone, C. W., and Colvin, Carl. "How to Study," Journal of Educational Psychology, XI (September, 1920), 348-54.
- Strang, Ruth. "Another Attempt to Teach How to Study," School and Society, XXVIII (October 13, 1928), 461-66.
- Taylor, Orville E. An Analysis of the Habits of Study of Superior High-School Students. Master's Thesis, Department of Education, University of Chicago, 1925.
- Thompson, William H. "An Experiment in Remedial Reading," School and Society, XXXIV (August 1, 1931), 156-58.
- Thorndike, E. L. "Reading as Reasoning," Journal of Educational Psychology, VIII (June, 1917), 323-32.
- Thorndike, E. L. "The Measurement of Achievement in Reading: Word Knowledge," Teachers College Record, XVII (November, 1916), 430-54.
- Tinker, Miles A. "The Role of Eye-Movements in Diagnostic and Remedial Reading," School and Society, XXXIX (February 3, 1934), 147-48.
- Traxler, Arthur E. The Measurement and Improvement of Silent Reading at the Junior-High School Level. Private Edition. Chicago: University of Chicago Libraries, 1932. Pp. 217.
- Watson, G., and Newcomb, T. M. "Improving Reading Ability Among Teachers College Students," Teachers College Record, XXXI (March, 1930), 535-39.
- Witty, P. A., and La Brant, Lou L. "Vocabulary and Reading," School and Society, XXXI (February 22, 1930), 268-72.
- Witty, P. A., and Lehman, H. C. "Teaching the College Student How to Study," Education, XLVIII (September, 1927), 47-56.
- Yoakum, G. A. Reading and Study. New York: The Macmillan Company, 1928.
- Zeleny, Florence K. "Remedial Instruction in Reading at the Freshman Level in a Teachers College," Educational Administration and Supervision, VIII (November, 1932), 607-20.
- Zimmerman, Regina. "Working Plan of Teaching How to Study," Educational Review, LXXIV (October, 1927), 168-71.

APPENDIX

The Appendix contains one form of the battery of reading tests devised in this study. All four forms of the test battery appear in the complete unpublished form of this dissertation which is filed in the Library of the University of Chicago.

A TEST OF ACHIEVEMENT IN SILENT READING

Form 1

Please fill the following blanks.

Name Age in years

High School Attended

City Date
(This city; where the test is taken)

Grade Boy Girl
(H. S. Senior, College Fr., etc.)

(Do Not Turn This Page Until Told to Do So)

DIRECTIONS. - This is a test to see how well you can read. (Follow all directions carefully.) On the following page there are some questions. (Do not turn to them now.) These questions do not ask for details but you cannot answer them unless you have read, and understood, the test paragraphs. You should read as rapidly as you can read UNDERSTANDINGLY. When the examiner gives the signal we will turn this page and read together the material at the bottom of Page 2. The examiner will read orally, and you are to read silently. When we come to the last word on Page 2 the examiner will stop reading, and you will simply turn the page and continue to read. After a while the examiner will say "mark". Put a circle around the word you are reading at that moment and keep right on reading. When you finish reading the test paragraphs you will find some printed directions telling you what to do next. YOU ARE NOT TO READ ANY PART OF THE TEST MATERIAL MORE THAN ONCE. Is there any question?

Read the directions above as many times as you like, but wait for the signal to turn to Page 2. When we do turn to Page 2, do not worry because the questions are printed upside down. They are printed that way purposely. The material which we will read together is at the bottom of Page 2.

* * * * *

Score A. . . .

Score C. . . .

Score B. . . .

- 1.- This essay shows that the elements necessary for success are (1) fact and common sense (2) mental alertness and good health (3) discernment and forethought (4) ambition and self-confidence (5) courage and perseverance..... (2)
- 2.- The author's primary purpose is to (1) define success (2) justify certain tendencies in the modern world (3) emphasize certain present-day trends in the field of government and religion (4) analyze certain factors which contribute to success (5) show that success is hard to achieve..... (1)
- 3.- An adequate insight into one's environment is said to result in (1) the ability to reason accurately (2) openmindedness (3) firm convictions or beliefs (4) an appreciation of the beauties of nature (5) a preference for the better things in life..... (1)
- 4.- The essay emphasizes the importance of a thorough knowledge of (1) scientific principles (2) political institutions (3) human nature (4) social principles (5) economic laws..... (1)
- 5.- It is stated that in order to win the confidence and esteem of men one must (1) have a character above reproach (2) see their point of view (3) be able to approach people tactfully (4) agree with them (5) perform some reason-ably valuable service for mankind..... (1)
- 6.- It is shown that the prevailing trends or currents of the age are to (1) oppose vigorously (2) modify (3) ignore entirely (4) comprehend clearly (5) followed consistently..... (1)
- 7.- One of the prevailing trends which is said to characterize the present century is (1) industrialization (2) scientific achievement (3) lawlessness (4) individualism (5) combination..... (1)
- 8.- In the field of government it is pointed out that the trend is toward (1) imperialism (2) democracy (3) communism (4) nationalism (5) political individualism..... (1)
- 9.- The supreme and final authority in religion is said to be (1) the church creed (2) the Bible (3) the Pope (4) scientific fact (5) one's conscience..... (1)
- 10.- Knowledge of the past needs to be supplemented by (1) discovery and invention (2) discernment of the future (3) scientific research (4) devotion to the present (5) idealism..... (1)
- Example.- The opposite of courage is (1) joy (2) fear (3) hate (4) pride (5) hope..... (2)

DIRECTIONS.- Draw a line under the word or phrase which should be used to complete each statement, and place the number of that word or phrase in the parentheses at the right. DO NOT TURN BACK TO THE TEST.

QUESTIONS ON THE TEST

Page 2

Page 2

AN ESSAY ON SUCCESS

If you were asked to explain the word life could you do it? Probably not; although you know well enough what the word life means. So it is with many other common ideas. We understand them, but only more or less vaguely. Let us think together about a very common idea, but one which is rich in meaning; namely, the idea expressed by the word "success". We shall not attempt to define the word

success. Every individual must define it for himself. But whatever the standard which one adopts, certain things may be safely pointed out as the necessary elements in its achievement. It is our task, therefore, to consider a few of these necessary elements and to show how success may be attained.

There are two forms of knowledge which the college can do but little to furnish, but which one must acquire if he is to succeed in life. Perhaps we should have said two forms of capacity, or, it may be that you would call them mental traits or characteristics. At any rate, what we have in mind can best be represented by the words insight and foresight. These two elements alone are probably not sufficient for the attainment of maximum success, yet they are so important that without them no other combination of traits can make success possible. As the most perfect electric motor stands idle and helpless except when it is supplied with electrical energy, so every individual must be impelled to success by the power of these two traits.

By the term insight we mean merely an adequate understanding of the fundamental principles and facts of life. A person with insight is one who knows how the world is put together. He understands how things work, and what it is that causes them to work the way they do. He sees the relation between cause and effect and is constantly looking for the why of every phenomenon. He thinks logically and reasons clearly. On the other hand, a person without insight is easily deceived. He does not insist on knowing why things behave the way they do. He falls an easy prey to the doctrines of superstition, prejudice, and quackery. His passive acceptance of the world is childlike in its simplicity, and he displays an attitude of naive credulity. His beliefs are without foundation and his conclusions are often irrational.

There are many types of insight which every intelligent person should acquire. One should seek to broaden his interests and should strive to get a practical knowledge of as many things as possible. For example, we do not all need to be electricians, but every person needs a practical knowledge of the fundamental principles of electricity, in order to be able to use the many electrical appliances that are found in the modern home. It is the same sort of practical, though incomplete knowledge that enables one to think logically and reason clearly; to become an intelligent citizen in our democracy. But more important than any other form of insight is a thorough understanding of people. In order to deal successfully with men one must understand them. He must know their natures and the motives which control their actions. In order to live peaceably among men, one must win their confidence, esteem, and respect. Mere eloquence and flattery will not suffice. In order to do this he must enter into their lives, see things as they see them, get their point of view, and must be able to put himself in their place. He must be able to do this not only for right-thinking, right-willing, and reasonable men, but he must be able to do it for wrong-thinking, prejudiced, and unreasonable men, because, for the most part, it is with such men that he will have to deal.

The other capacity which one should develop is foresight. There is an unavoidable destiny which shapes the course of human affairs. One should observe what its forces are doing, and to what end it is shaping our generation. We are like men upon the deck of an ocean steamer. We can walk forward or backward or not walk at all, but we are borne on toward a distant point by forces which transcend our understanding. The currents which control our lives are almost wholly beyond our control, and success depends upon the extent to which we perceive what those currents are and whither they are carrying us. We incur failure, however sagacious, strong, talented, and learned, if we set ourselves against those life currents endeavoring to thwart them, or live in ignorance of them as though they did not exist. Let us illustrate by citing certain trends.

The individualism of the middle of the nineteenth century belongs to the past. This century is an era of combination. Ruthless and unlimited competition is

passing away and no human forces are strong enough to restore it. The man who endeavors to prevent combination whether of capital or of labor, swims against a resistless tide. The wise leader of industry will neither ignore the competition of the past nor attempt to resist the tendency toward co-operation and combination in the present. He will endeavor to see how, out of the competition of the past, to construct a combination and co-operation which will preserve the virtues and values of individualism, and secure also the virtues and values of a true, co-operative fellowship.

The tendency toward combination is as apparent in government as it is in industry. It is a part of that great movement toward universal co-operation which can be guided but cannot be thwarted. The separate and jealous provinces of Italy are united into one Italian kingdom. The hostile German states now constitute a single German empire. Egypt, India, and Australia have become parts of the great British Empire. The frantic endeavor for the dissolution of the American commonwealth in the interests of political individualism -- the Civil War -- failed. In lieu of it, the American flag floats not only over all the former commonwealth, but in spite of the protests of the timid, carries American sovereignty across the Pacific to Hawaii and the Philippines. Commerce beckons American enterprise to enter the Orient, which lies a thousand miles nearer to our coast than to the British Isles. The opportunities of the present demand of us a courage and a wisdom equal to the exigencies of the new day which dawned with the twentieth century. To resist this tendency toward what is called imperialism is a hopeless task; to guide that tendency so that wherever the American flag carries with it American sovereignty it shall also carry with it liberty, justice, and goodwill, is the problem which the people of this generation must learn to solve.

Another trend may be cited from the field of religion. The supreme and final authority has passed for substantially all Protestant denominations from the Church, and is passing, despite the hopeless resistance of many, from the Bible. The minister of the future must learn that the secret of authority is God in the individual conscience, and that the authority of the Church and of the Bible is due to the fact that they have both interpreted that voice to human souls. He who would succeed in the religious development of his generation must not live in the nineteenth century; still less in the sixteenth. He must live in the twentieth century and must show men how to use both the Church and the Bible in interpreting that voice of God in the soul of man, which is the secret of all authority -- social, political, and religious.

You may not agree with this interpretation of the life currents of this twentieth century. We briefly describe them here, not for the purpose of making you see life as we see it, but for the purpose of inducing you to turn your faces toward the future and see for yourselves whither those life currents are guiding you. The knowledge of the past will not give you success in the present, unless out of that past you learn lessons to guide you toward the future. All that you may learn of history and literature, of science and of art, in your college course, will serve to enrich your life and that of your fellow men, will serve to give you real influence over them and real leadership in your age, only, as to the scholarship which the college has conferred upon you, you add by your own individual study of life and of men those two vital qualities -- insight and foresight. You must have an understanding of the world and especially of men, and some comprehension of the future toward which you must guide both them and yourself.

- - - - -

Name _____

Score _____

TEST OF READING VOCABULARY

Form A

DIRECTIONS.- In each of the exercises below you are to select the best definition for the word which is underlined in the sentence or phrase. Draw a line under the word or phrase thus selected, and place its number in the parentheses at the right. Here is an example:

EXAMPLE.- Let us try to appease him.

(1)allure (2)deceive (3)pacify (4)believe-in (5)convince (3)

1.- His acute interpretation.

(1)pleasing (2)appropriate (3)impressive (4)penetrating (5)idealistic ()

2.- The most amicable relationship.

(1)advantageous (2)honorable (3)friendly (4)delightful (5)enduring ()

3.- The apprehension of the man.

(1)pursuit (2)seizure (3)indictment (4)trial (5)condemnation ()

4.- Why augment the faults of others?

(1)condemn (2)consider (3)point-out (4)magnify (5)attempt-to-remedy ()

5.- A marked aversion.

(1)dislike (2)contrast (3)distrust (4)copy (5)peculiarity ()

6.- He spoke to the baronet.

(1)titled-gentleman (2)wife-of-the-baron (3)lawyer (4)judge (5)clergyman ()

7.- Belligerent nations.

(1)powerful (2)allied (3)hostile (4)imperialistic (5)prosperous ()

8.- His leading interests are in carnal affairs.

(1)social (2)moral (3)athletic (4)intellectual (5)unspiritual ()

9.- Is he not your co-equal?

(1)peer (2)rival (3)friend (4)former-classmate (5)business-partner ()

10.- A commodious house.

(1)gorgeous (2)cozy (3)vine-covered (4)sheltered (5)spacious ()

11.- Their opinions concur.

(1)harmonize (2)mislead (3)change (4)suffice (5)clash ()

12.- A special convocation.

(1)invitation (2)assembly (3)announcement (4)arrangement (5)offer ()

13.- The boy is too credulous.

(1)studious (2)wasteful (3)easily-deceived (4)reckless (5)critical ()

14.- His cupidity made him unpopular.

(1)egotism (2)frankness (3)greed (4)ill-manners (5)gossip ()

15.- A diffident pupil.

(1)timid (2)lazy (3)peculiar (4)mischievous (5)stubborn ()

F. A

16.- A discursive treatment of the topic.
(1)brief (2)rambling (3)satisfactory (4)humorous (5)coherent ()

17.- It will enhance his opportunity.
(1)threaten (2)diminish (3)end (4)insure (5)improve ()

18.- Fallow ground.
(1)sacred (2)rough (3)uncultivated (4)worthless (5)barren ()

19.- He pointed out the gargoyle.
(1)kind-of-fish (2)monastery (3)ancient-ship (4)carved-water-spout (5)coat-of-arms ()

20.- A graphical account.
(1)vivid (2)written (3)fantastic (4)funny (5)prejudiced ()

21.- Grist from the mill.
(1)chaff (2)screenings (3)by-products (4)meal (5)dust ()

22.- A heterogeneous mass.
(1)seething (2)conglomerate (3)loathsome (4)compact (5)colorful ()

23.- Living in ignominy.
(1)poverty (2)fear (3)seclusion (4)dishonor (5)obscurity ()

24.- Imminent disaster.
(1)threatening (2)wide-spread (3)certain (4)resulting (5)terrible ()

25.- Infusion of the doctrine.
(1)interpretation (2)decline (3)criticism (4)rejection (5)introduction ()

26.- The conclusions are irreconcilable.
(1)valid (2)inconsistent (3)significant (4)obvious (5)premature ()

27.- A jaunty appearance.
(1)rustic (2)neat (3)gay (4)ludicrous (5)tourist-like ()

28.- A lactic substance
(1)sparkling (2)milky (3)sweet (4)sticky (5)luke-warm ()

29.- A spirit of levity.
(1)vengeance (2)reverence (3)tolerance (4)frivolity (5)unrest ()

30.- A lucid description.
(1)droll (2)brief (3)fantastic (4)poetic (5)clear ()

31.- He spoke as to a minion.
(1)slave (2)child (3)dog (4)tramp (5)very-dear-friend ()

32.- Evidence which will mitigate his punishment.
(1)determine (2)diminish (3)postpone (4)remove (5)prolong ()

33.- Queer mummery.
(1)embalming (2)mausoleums (3)designing (4)pronunciation (5)masquerading ()

34.- A pandemoniac performance.
(1)tragic (2)gay (3)silent (4)riotous (5)memorial ()

F. A

- 35.- He became quite a pedant.
 (1)nuisance (2)vain-scholar (3)tourist (4)pessimist (5)philosopher ()
- 36.- Accused of perjury.
 (1)theft (2)contempt-of-court (3)bribery (4>false-swearing (5)fraud ()
- 37.- Portentous signs.
 (1)numerous (2)foreboding (3)never-failing (4)old-Indian (5)faint ()
- 38.- A satisfactory redress.
 (1)solution (2)report (3)state-of-affairs (4)compensation (5)investigation ()
- 39.- He can regale them properly.
 (1)entertain (2)repay (3)reward (4)advise (5)dress ()
- 40.- The remission of his claim.
 (1)payment (2)proof (3)recognition (4)recording (5)surrender ()
- 41.- He will gladly requite you.
 (1)reward (2)release (3)forgive (4)assist (5)inform ()
- 42.- A salient fact.
 (1)controversial (2)recently-established (3)scientific (4)discredited
 (5)significant ()
- 43.- The growth of sedition.
 (1)socialism (2)depression (3)dissension (4)unemployment (5)crime ()
- 44.- A clever sleight.
 (1)plan (2)trick (3)remark (4)invention (5)combination ()
- 45.- Spurious rumors.
 (1)many (2)dangerous (3)injurious (4)rapidly-spreading (5)unauthentic ()
- 46.- The subsidy makes it possible.
 (1)law (2)amendment (3)gift (4)court's-interpretation (5)charter ()
- 47.- They surfeit us with their flattery.
 (1)amuse (2)delight (3)injure (4)disgust (5)insult ()
- 48.- Turbid water.
 (1)rushing (2)bitter-tasting (3)muddy (4)distilled (5)boiling ()
- 49.- The plan was too unartful.
 (1)complicated (2)fantastic (3)obscure (4)unfair (5)simple ()
- 50.- Usurpation of power.
 (1)abuse (2)exercise (3)unlawful-assumption (4)limitation (5)relinquish-
 ment ()

Name _____

Score _____

A TEST OF ACCURATE INTERPRETATION IN READING
Form 100

DIRECTIONS.— This is a test to see how accurately you can interpret what you read. You will first read a paragraph and then answer questions about it. In order to find out what you are to do, read the following paragraph and answer the questions about it.

SAMPLE PARAGRAPH.— Advertising has changed remarkably in recent years. Its methods used to be quite simple and crude as compared with the highly organized departments and carefully planned advertising campaigns which feature modern business. For example, I remember the leading merchant in a little town out West. His store stood on the town's busiest corner, where his three sons helped him sell drygoods. Every Saturday noon in winter he used to get up on the roof of his store and throw a new overcoat into the street. The crowd scrambled for it. That was advertising in those days. Incidentally, one of his sons was supposed to be quick enough to grab the coat and rush it back to the shelves to be thrown down again next Saturday. It was a bad day for the boys if they let an outsider get away with that decoy coat!

There will be three kinds of questions about each paragraph. The first will require an understanding of the total meaning of the paragraph. Try the following sample question. Select the best answer; underline it; and place its number in the parentheses at the right.

- I. The western merchant is portrayed as (1)opposed to advertising (2)a miserly money grabber (3)a typical advertiser of his day (4)a man with modern business ideas (5)a harsh father. ()

The correct answer to the question above is number "3". Underline the phrase "a typical advertiser of his day" and place a figure "3" in the parentheses at the end of the dotted line.

Now try the next question. Three of the following phrases can be used to make a partial outline of the paragraph. Here is the exercise:

II. Make a partial outline using three of the following phrases:

- | | |
|--|-----------------|
| (1)Saturday noons. | I. () |
| (2)Grab the coat | 1 () |
| (3)A decoy coat. | a () |
| (4)Back to the shelves | |
| (5)Advertising methods | |
| (6)Former methods simple and crude | |

Number "5" is the most important idea among them (the very topic of the paragraph). Underline it; and place a figure "5" in the parentheses after the "I" in the schematic outline at the right.

Statement number "6" grows out of number "5" in that it explains what is said about advertising methods. Therefore, you should underline it (do it), and place a figure "6" after the "1", above.

Number "3" suggests the illustration used to show that number "6" is true. Therefore, you should underline it (do it), and place a figure "3" after the letter "a" in the schematic outline above.

(Directions continued on next page)

In the foregoing exercise, do you see why phrases "5", "6", and "3" make a partial outline of the paragraph? It is because they represent important ideas, and because they are so closely related that each leads logically to the next or grows out of the one which precedes it. Now try the next exercise.

III. Indicate whether these statements are true, false, or unrelated: (See directions below.)

- 1.- Advertising methods have been changed considerably. ()
- 2.- Modern advertising is highly organized. ()
- 3.- The western merchant used modern advertising methods. ()
- 4.- The merchant gave away a coat each Saturday in the year ()
- 5.- Giving away merchandise is not good advertising ()
- 6.- When the coat was thrown into the street one of the merchant's sons tried to recover it. ()

You are to put a plus sign (+) after the true statements, a minus sign after those which are false (-), and a zero (0) after those which are unrelated. Therefore, in the exercise above you should put plus signs after statements "1", "2", and "6". You should put a minus sign after statements "3" and "4", and a zero after number "5". (Do it) NOTE WELL.- A statement is to be considered true if the paragraph makes it true, and false if the paragraph makes it false. It is unrelated if the paragraph says nothing about it which would allow you to draw a reasonable conclusion. Remember -- what you know, or what you think, has nothing to do with the matter. The paragraph MAKES it true; MAKES it false; or MAKES it unrelated.

Each paragraph in the test will have these three types of exercises. You are to read (1) in order to get the total meaning, (2) in order to organize the material, and (3) in order to answer fact questions. You may re-read any paragraph as often as you like, but do not waste time in needless re-reading. You should be able to finish the test if you work at a reasonable rate, but you will have no time to waste. Refer back to these directions if you need to do so. Do you have any question about the directions? Then, continue with the test paragraphs which follow.

TEST PARAGRAPHS

101

It was a dark night, just after midnight, and everything was as still as death except for the washing of the water by the vessel's side. We were going along at a great rate, dead before the wind, and the little brig, covered with canvas, was doing great business with very little noise. Most of the crew were below; even all our watch except myself and the man at the wheel were asleep under the lee of the boat. I was walking to and fro on the deck to keep awake, when suddenly, we heard a loud scream which seemed to come directly from under the bows of the ship. The darkness, the stillness of the night, and the solitude of the ocean gave to the sound a dreadful and almost supernatural effect. I stood perfectly still, and my heart beat rapidly. The sound woke up the rest of the watch who stood looking at each other. My first thought was that we had run down a small boat in the darkness; a boat carrying the crew of some wrecked vessel, or the boat of some whaleship. Another scream; but less loud than the first! This started us, and we ran forward and looked over the bows, over the sides, and to leeward. Not a thing was to be seen or heard. What should be done? Call the captain, and heave the ship aback? Just at this moment in crossing the forecastle one of the men saw a light below, and looking down the scuttle saw the watch all out of their berths, dragging, shaking, and scolding one of their fellows to awake him out of a nightmare. We made a joke of the matter, and well we might laugh, for we were not a little relieved by its ridiculous termination.

I. The most accurate title for this narrative would be (1)the sailor's dream
(2)a lonely watch (3)a rude awakening (4)a false alarm (5)a ridiculous situation
..... ()

II. Make a partial outline using three of the following phrases:

- | | |
|--------------------------------|-------------|
| (1)A quiet watch. } | I. () |
| (2)Only a nightmare } | 1. () |
| (3)Seeking the source } | a. () |
| (4)Walking to and fro } | |
| (5)A strange sound. } | |
| (6)The sleeping crew. } | |

III. Indicate whether these statements are true, false, or unrelated:

1. The outcry caused considerable consternation among the sailors. ()
2. The scream sounded more dreadful because everything was unusually quiet .. ()
3. Sailors are quite generally inclined to be superstitious. ()
4. The watch decided to call the captain and to turn the boat about. ()
5. The sailor who had had the nightmare was greatly embarrassed. ()
6. The sound seemed to come directly from behind the vessel. ()

One of the newer servants of man is the automobile. The important place which it occupies in the scheme of modern affairs is too well recognized to need elaboration. But with its many benefits to humanity has come a growing menace in the form of fatal accidents. While the death rate from small-pox, diphtheria, and tuberculosis has been decreasing steadily, the corresponding death rate from automobile accidents has been increasing. During the past decade it has almost doubled, until today it is approximately 20.5 per 100,000 of population. This alarming increase in fatalities has caused considerable alarm, and justly so. Certainly many of the accidents which have occurred might have been prevented by the exercise of reasonable precaution. Yet it is easy to overemphasize the seriousness of the problem because of the spectacular way in which the death toll is taken. As a matter of fact the potential danger incurred in automobile travel is no greater than that for many other types of transportation. The death toll is greater merely because more people are riding in automobiles. On the basis of miles travelled the automobile is safer than either the train or the aeroplane. On the basis of person-mile units the automobile is as safe as either. There is every reason why preventable accidents should be eliminated, but it is altogether unlikely that the death rate from automobile accidents will thereby be decreased. An appreciable gain in that direction cannot reasonably be expected until the point has been reached where there is no further increase in the use which is made of the automobile.

I. It is the author's purpose to show that (1)auto accidents are a necessary evil
(2)most accidents could be prevented (3)automobile travel is relatively safe (4)
auto accidents are likely to increase rather than decrease (5)the auto has brought
more than enough benefits to offset its injuries. ()

II. Make a partial outline using three of the following phrases:

- | | |
|--|-------------|
| (1)Widespread use of automobiles. } | I. () |
| (2)The scheme of modern affairs } | 1. () |
| (3)The dangers of automobile travel } | a. () |
| (4)Other types of transportation. } | |
| (5)Numerous accidents } | |
| (6)Exercising reasonable precaution } | |

III. Indicate whether these statements are true, false, or unrelated:

1. Automobile fatalities are disproportionately large. ()
2. Nearly all accidents are caused by unavoidable circumstances. ()
3. The problem of automobile accidents is not really a serious one. ()
4. Automobile fatalities attract more attention than many other types because they are sudden and more or less dramatic. ()
5. The death rate from auto accidents is higher than that from any other single cause. ()
6. The automobile is a relatively safe means of conveyance. ()

103

A maned lion rose and galloped off through the tall grass. Immediately a band of sixty native hunters gave chase. They were splendid savages, stark naked, lithe as panthers, the muscles rippling under their smooth dark skins. All their lives they had lived on animal food, milk, blood, and flesh, and they were fit for any fatigue or danger. Each carried an ox-hide shield and the formidable war spear, whose narrow head was burnished till it shone like silver, and whose point and edges were razor sharp. Each sinewy warrior carried his heavy weapon as if it were a toy, twirling it till it glinted in the sunlight. The faces of the hunters were proud, cruel, and fearless as they pursued the lion with long, springy strides. The pursuit was swift and relatively short, for the lion halted and turned under a low thorn tree scarcely a mile from where he had risen. He was a magnificent beast with a tawny black mane; in his prime, teeth and claws perfect, with mighty thews and savage heart. His life had been one unbroken career of rapine and violence. One by one the spearmen came up, at a run, and began to form a ring about him. He remembered other rings like this, and his anger mounted instantly. Crouching behind their shields, spear in hand, their eager faces peered over the rims of their shields toward the grim master of the jungle. First one way and then the other the lion turned, surveying the only foes who dared to molest him. His mane bristled, his tail lashed; he held his head low, the upper lip now drooping over the lower jaw, now drawn up so as to show the gleam of the long white fangs. Intent, silent, with shields held steady and quivering spears poised, the hunters braced themselves and waited. The great man-killing beast, intent on blood, gave one quick look, saw where the line was thinnest, vented his thunderous wrath in a murderous grunting roar, and crouched very low. The moment had come.

I. It is the writer's purpose to portray (1)an exciting chase (2)an exhibition of courage (3)a feat of valor (4)the murderous ferocity of a lion (5)methods of lion hunting used by natives. ()

II. Make a partial outline using three of the following phrases:

- | | | | |
|--------------------------------------|---|------------|-----|
| (1)The challenge is given. |) | I. | (.) |
| (2)A narrow escape |) | 1. | () |
| (3)Dauntless courage |) | a. | () |
| (4)The death of a killer |) | | |
| (5)Wreaking vengeance. |) | | |
| (6)Formidable enemies meet |) | | |

III. Indicate whether these statements are true, false, or unrelated:

1. The lion had been surrounded by natives before this encounter. ()
2. The natives were hunting this lion because of his recent raids. ()
3. Both the lion and the hunters seemed to be eager for the conflict. ()
4. The only protection which the hunters enjoyed was that of their spears. ()
5. The initiative in the final combat came from the lion. ()
6. The natives were experienced in this method of hunting lions. ()

Two mistaken attitudes may be assumed towards the traditions of man's civilization. They may be regarded as good merely because they are old, or as bad because of their antiquity. The total reactionary is, of course, as rare as the extreme radical. Obviously, the uncritical acceptance of tradition and the uncritical rejection of it are both to be avoided. Tradition is neither false nor true because it has been handed down from the past. The presumption is that there was some ground in the past for an idea or way of behaving, else it could not have become part of man's culture. What that ground may have been and whether it is still present is not always easy to determine. There was a sound basis for primitive man's prohibition of marriage between kindred, but it took modern biological science to make clear that he thereby avoided the transmission of feeble-mindedness to his offspring. Other traditions and prohibitions were based merely on primitive taboo and are unable to stand the criticism of modern science. The critical evaluation of tradition must take place as man's civilization is assimilated in each generation. Only in the light of an intelligent analysis of the situation can the question as to the retention of a tradition be properly answered. One cannot say that all tradition is true because some of it is, or that it is wholly false because part of it has been found so. There are relatively stable elements in civilization and there are other parts which are moving rapidly, as their revision goes on. Every young person should assimilate as quickly as possible the valuable core of tradition and then proceed in the field in which he or she specializes, to keep abreast of the changes which are taking place. The bulk of knowledge is so large, however, that it is the rare man today who gets to the real frontier of knowledge in any field. The mass of mankind must evaluate tradition in terms of knowledge which they have received from others.

I. The author's point of view is best expressed by saying that (1)each generation should evaluate the traditions which it receives (2)tradition has not been able to stand the tests of science (3)tradition is the foundation upon which progress must be built (4)every tradition is partly true and partly false (5)the most stable traditions need frequent revision. ()

II. Make a partial outline using three of the following phrases:

- | | |
|---|----------------|
| (1)Neither inherently good nor bad. | I. () |
| (2)Each generation of mankind | 1. () |
| (3)Constant evaluation. | a. () |
| (4)Traditions in our culture. | |
| (5)Keeping up with the times. | |
| (6)Quick assimilation | |

III. Indicate whether these statements are true, false, or unrelated:

1. The value of traditions which are followed more or less blindly at first is sometimes proven by the later findings of science. ()
2. Traditions should be either wholly accepted or wholly rejected ()
3. There is some justifiable ground for every tradition. ()
4. Most people become pioneer thinkers in some one field of endeavor. ()
5. Feeble-mindedness is transmitted only by parents who are related. ()
6. Traditions which were justifiable in one generation may not be appropriate in the next generation ()

There is one principal and as it were radical distinction between different minds with respect to philosophy and the sciences, which is this: some minds are stronger and more apt to mark the differences between things, others to mark their resemblances. The steady and acute mind can fix its contemplations and dwell and fasten on the subtlest distinctions. The lofty and discursive mind, on the other hand, recognizes and puts together the most general resemblances. The attainment of either a respectable science or a consistent philosophy demands the employment of both types of thinking. That the same individual so rarely succeeds in discerning both the similitude and dissimilarity in the objects of his study means that the progress of man in his conquest of truth must be slow. The explanation of this mental phenomenon is probably to be found in the fact that one's attention span is naturally somewhat limited. When the mind is intent upon one type of observation it is practically immune to stimulation of every other sort. Moreover, a division of the attention is so detrimental to the keenness or alertness of the mind that where the attention is divided flagrant oversight is almost certain to result. It is better that the penetrating discriminations of the one type of mind be supplemented with the coherent combinations of the other, than that mediocrity should prevail in both fields. We must have both the analyses of the former and the syntheses of the latter, even though both frequently err in excess; the one by catching at gradations, the other at shadows.

I. The paragraph is written to demonstrate that (1)everyone is fitted to do some one thing best (2)the range of attention is limited (3)the development of a science is a slow, laborious task (4)science is the result of co-operative enterprise (5)one is seldom keen in making both comparisons and contrasts. ()

II. Make a partial outline using three of the following phrases:

- | | |
|---|----------------|
| (1)The span of attention. | I. () |
| (2)Leading to serious oversights. | 1. () |
| (3)Supplementary functions. | a. () |
| (4)Types of minds | |
| (5)A demand for both types. | |
| (6)Immune to outside stimulation. | |

III. Indicate whether these statements are true, false, or unrelated:

1. The attempt to become equally skillful in noting similarities and differences is likely to result in mediocre performance. ()
2. Personal will power is the only limit to the attention span. ()
3. It is an easier form of mental activity to note likenesses than to distinguish differences. ()
4. The most alert observers are often guilty of flagrant oversight. ()
5. It is improbable that man's conquest of truth will ever be rapid. ()
6. A true science or philosophy can never be formulated until each worker becomes equally proficient in both types of thinking. ()

